UNIVERSIDAD DE LA RIOJA

TRABAJO FIN DE ESTUDIOS

Título

Introducción a la resolución de sistemas de ecuaciones polinómicas

Autor/es

ADRIÁN LATORRE RODRIGO

Director/es

JESÚS ANTONIO LALIENA CLEMENTE

Facultad

Facultad de Ciencia y Tecnología

Titulación

Grado en Matemáticas

Departamento

MATEMÁTICAS Y COMPUTACIÓN

Curso académico

2016-17

publicaciones.unirioja.es
E-mail: publicaciones@unirioja.es

UNIVERSIDAD DE LA RIOJA

Facultad de Ciencia y Tecnología

TRABAJO FIN DE GRADO

Grado en Matemáticas

Introducción a la resolución de sistemas de ecuaciones polinómicas

Alumno:

Adrián Latorre Rodrigo

Tutores:

Jesús Laliena Clemente

Logroño, febrero, 2017.

Índice

Resumen.

Este trabajo es un curso introductorio a la resolución de sistemas de ecuaciones de polinomios. Esta memoria es de carácter teórico-práctico, por lo que en su presentación vamos a ir desarrollando numerosos ejemplos a la vez que explicamos la Teoría de Bases de Groebner, que es la herramienta que utilizaremos para intentar resolver estos sistemas.

Finalmente veremos una aplicación de esta teoría en Robótica.

Abtract.

This project covers an introductory course on the resolution of systems of polynomial equations. In this theoretical and practical report, different systems of diverse types will be illustrated with many examples, while we are going to describe Groebner's Basis theory, tool used to solve them. Finally we will see an application of this theory.

1. Introducción.

Supongamos que queremos resolver el siguiente problema: calcular los máximos y mínimos de la función $f(x, y, z) = x^3 + 2xyz - z^2$ sujeta a la restricción $g(x, y, z) : x^2 + y^2 + z^2 = 1$. El método de los multiplicadores de Lagrange nos dice que $\triangle f = \lambda \triangle g$ en un mínimo o máximo local. Esto nos da el siguiente sistema de ecuaciones para las variables x, y, z, λ

$$\left.\begin{array}{lcl} 3x^2 + 2yz & = & 2x\lambda \\ 2xz & = & 2y\lambda \\ 2xy - 2z & = & 2z\lambda \\ x^2 + y^2 + z^2 & = & 1. \end{array}\right\}$$

El ejemplo que acabamos de ver nos ha llevado a un sistema de ecuaciones polinómicas. En la vida real infinitud de problemas desembocan en un sistema de ecuaciones de polinomios. Estos sistemas pueden ser de dos tipos:

(i) Sistemas de ecuaciones lineales,

(ii) Sistemas de ecuaciones polinómicas no todas lineales.

La resolución de sistemas del segundo tipo no es un tema trivial. Incluso en algunos casos no podemos llegar a conocer las soluciones exactas debido a la gran complejidad que pueden llegar a alcanzar sus soluciones.

En cualquier caso, el estudio de su solución es imprescindible a la hora de abordar la resolución de problemas de la vida cotidiana dentro de muy diversos campos: ingeniería, biología, arquitectura, economía, telecomunicaciones, transportes, etc...

En los últimos años nuestra habilidad para manipular sistemas de ecuaciones expresadas mediante polinomios ha experimentado algunas transformaciones cruciales. Las bases de Groebner se introdujeron en 1965, junto con un algoritmo para calcularlas (el algoritmo de Buchberger).

Bruno Buchberger nació el 22 de octubre de 1942 en Innsbruck. Es profesor de Matemáticas en computación para la Universidad Johannes Kepler de Linz (Austria). En 1965, en su Ph.D. tesis, creó la teoría de bases de Groebner. A lo largo de su carrera ha ido desarrollando esta teoría (ver por ejemplo [3],[4]). Llamó a estos objetos con el mismo nombre que su tutor de tesis Wolfgang Groebner.

En 2007, Buchberger recibió el Premio de Teoría y Práctica Paris Kanellakis concedido por la ACM (Association for Computing Maachinery) por su trabajo sobre las bases de Groebner. Es necesario comentar que, el matemático ruso N.M. Gjunter había utilizado un concepto similar en 1913, publicado en diversas revistas matemáticas rusas.

Actualmente este estudio, apoyado por el espectacular crecimiento de las capacidades de los ordenadores modernos y las muchas herramientas desarrolladas por la geometría algebraica clásica, ha ganado una gran importancia.

La teoría de las bases de Groebner se ha investigado por muchos autores en varias direcciones y se ha generalizado a otras estructuras como los polinomios sobre anillos de ideales principales. Recientemente, las bases de Groebner han sido aplicadas a multitud de problemas por su capacidad de resolver sistemas de ecuaciones polinómicos y como modelo algebraico de computación.

En este trabajo veremos lo siguiente:

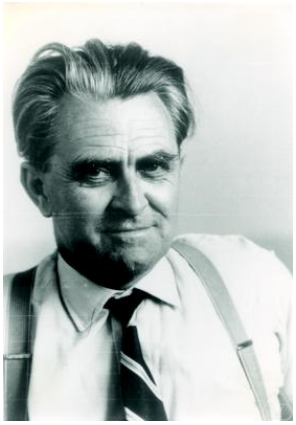


Figura 1: Wolfang Groebner (1899–1980) fue un matemático austriaco-italiano que destacó en temas relacionados con la geometría algebraica y álgebra computacional. Fue el tutor de tesis de B.Buchberger.

(i) En la sección 2 introduciremos las herramientas necesarias para comprender la teoría que vamos a desarrollar.

(ii) En la sección 3 definiremos un algoritmo de división para polinomios en varias variables y también las bases de Groebner, y veremos un modo de calcularlas.

(iii) En la sección 4 (la más difícil de entender) veremos que el algoritmo de cálculo de las bases de Groebner es en realidad una generalización de la eliminación Gaussiana en sistemas de ecuaciones lineales. Este algoritmo cambia un sistema de ecuaciones polinómicas por otro equivalente en el que las ecuaciones van teniendo cada vez menos variables. El Teorema de Eliminación nos muestra esto, y el de Extensión nos dice cuándo se pueden extender las soluciones parciales de los polinomios con menos variables a soluciones del sistema. En esta sección trataremos también el problema inverso al de resolver sistemas de ecuaciones polinómicas, esto es, a partir de un conjunto de puntos dados de forma paramétrica por funciones polinómicas en varias variables, se trata de encontrar las ecuaciones implícitas que satisfacen los mismos. Este problema se resuelve también con bases de Groebner.

(iv) En la seccion 5, finalmente, veremos algunas aplicaciones de las bases de Groebner al campo de la Robótica.

El uso de las bases de Groebner es algo que actualmente se sigue estudiando. No siempre podremos resolver los sistemas, y dependiendo del orden monomial que usemos podremos llegar más o menos fácilmente a la solución. Incluso, habrá momentos que no podamos llegar a una solución, y entonces habrá que pensar en la posible ayuda de métodos numéricos.

Esta introducción así como la memoria están basados en el libro *Ideals, Varieties,and Algorithms* escrito por David Cox, Jonh Little y Donald OShea (ver [5]). Otros libros sobre el tema o sobre la teoría de variedades afines son por ejemplo [1],[2],[6] y [7]. En el desarrollo de este trabajo se han omitido demostraciones de algunos resultados. Por falta de espacio, sólo aparecerán fundamentalmente las demostraciones de los resultados más importantes.

Por último, añadir que para la correcta comprensión de esta memoria, es necesario resultados más o menos básicos sobre álgebra lineal así como de geometría.

2. Espacios y Variedades Afines. Polinomios en una variable.

En esta sección introduciremos los conceptos básicos de la teoría que desarrollaremos en el trabajo. Estos conceptos son de tipo algebraico geométrico y referentes a polinomios en varias variables, variedades afines, ideales de polinomios en anillos de varias variables y polinomios en una variable.

A lo largo de estas páginas usaremos con frecuencia la terminología que definiremos a continuación:

Definición 2.1. *Sea el polinomio en las variables $x_1, ..., x_n$ con coeficientes en K que denotaremos:*

$$f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \, ; \qquad a_\alpha \in K$$

(i) *Cada* monomio *es de la forma* $x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ *con*

$$\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n.$$

(ii) *Llamamos* a_α *al* coeficiente *del monomio* x^α.

(iii) *Si* $a_\alpha \neq 0$ *llamamos* término de f *a* $a_\alpha x^\alpha$.

(iv) *Llamaremos* grado del monomio *a* $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

(v) *Llamaremos* grado de f *al* máx$\{|\alpha| \ : \ a_\alpha \neq 0\}$.

(vi) *Al conjunto de todos los polinomios lo denotaremos* $K[x_1, ..., x_n]$.

Definición 2.2. *Dado un cuerpo K y $n \in \mathbb{N}$ definimos* espacio afín n-dimensional *sobre K como el conjunto:*

$$K^n = \{(a_1, ..., a_n) \ : \ a_1, ..., a_n \in K\}$$

Definición 2.3. *Sea K un cuerpo y $f_1, ..., f_s$ polinomios en $K[x_1, ..., x_n]$. LLamamos* variedad afín *al conjunto:*

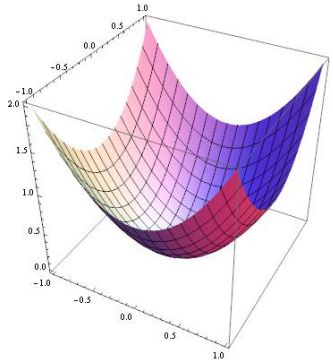
$$\mathbb{V}(f_1, ..., f_s) = \{(a_1, ..., a_n) \in K^n \ : \ f_i(a_1, ..., a_n) = 0 \quad \forall \, 1 \leq i \leq s\}$$

Ejemplo 2.4. *Vamos a resolver la siguiente ecuación:*

$$x^2 + y = z.$$

Con lo visto hasta ahora, sabemos que las soluciones de esta ecuación son las raíces del polinomio $f = x^2 + y - z \ \in \mathbb{R}[x, y, z]$, que se corresponden con la variedad afín $\mathbb{V}(f)$.

Si dibujamos todas las raíces de dicho polinomio f obtenemos un paraboloide de revolución (que viene dado por la siguiente gráfica):

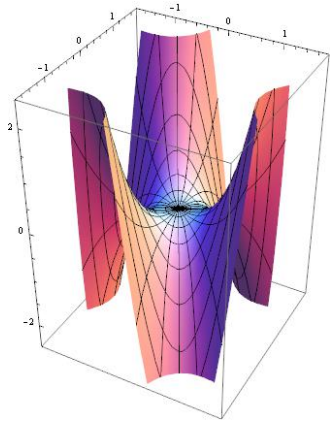


Ejemplo 2.5. *Resolveremos la siguiente ecuación:*

$$z = x(x^2 - 3y^2)$$

Al igual que en el ejemplo anterior, las soluciones de esta ecuación se corresponden con las raíces del polinomio $f = x(x^2 - 3y^2) - z \in \mathbb{R}[x, y, z]$. *Al conjunto de todas estas raíces es a lo que llamamos la* variedad afín determinada por f, $\mathbb{V}(f)$. *Es decir que todos los puntos de la variedad son las raíces del polinomio o, lo que es lo mismo, las soluciones de nuestra ecuación inicial.*

Si pintamos la variedad obtenemos la siguiente superficie:



Ejemplo 2.6. *Resolveremos la siguiente ecuación:*

$$r = \text{sen}(2\sigma)$$

Ahora no tenemos una superficie, sino una curva en el plano dada en forma polar. Esta curva es llamada Rosa de cuatro hojas. *Dicha curva se corresponde con la variedad afín* $\mathbb{V}((x^2 + y^2)^3 - 4x^2y^2)$ *en el plano afín* $\mathbb{R}^2$.

Vamos a demostrar que es cierto, probándolo por doble contenido:

(i) *Usando* $r^2 = x^2 + y^2$ *siendo* $x = r\cos(\sigma)$ $y = r\,\mathrm{sen}(\sigma)$, *vamos a ver que cualquier punto de la* Rosa de cuatro hojas *está en* $\mathbb{V}((x^2+y^2)^3 - 4x^2y^2)$.

Al ser una variedad formada por un único polinomio sabemos que los puntos de la variedad tienen que cumplir que $(x^2+y^2)^3 - 4x^2y^2 = 0$. *Si sustituimos los x, y que teníamos antes, y usando que* $r = \mathrm{sen}\,2\sigma$, *llegamos a:*

$$[[r\cos(\sigma)]^2 + [r\,\mathrm{sen}(\sigma)]^2]^3 - 4[r\cos(\sigma)]^2[r\,\mathrm{sen}(\sigma)]^2 = 0.$$

(ii) *Nuestro objetivo es ver que todo punto de* $\mathbb{V}((x^2+y^2)^3 - 4x^2y^2)$ *pertenece a la rosa de 4 hojas, es decir debemos probar*

$$\{(x,y) : (x^2+y^2)^3 - 4x^2y^2 = 0\} \subseteq \{(\mathrm{sen}(2\sigma)\cos\sigma, \mathrm{sen}(\sigma)) : \sigma \in [0, 2\pi)\}$$

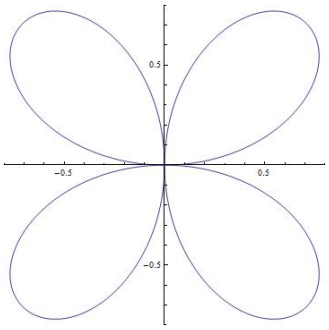
Comencemos operando y llegaremos a la rosa de 4 vientos:

$$(r^2\cos^2(\sigma) + r^2\,\mathrm{sen}^2(\sigma))^3 - 4r^2\cos^2(\sigma)r^2\,\mathrm{sen}^2(\sigma) = 0$$

$$r^4(r^2 - 4\cos^2(\sigma)\,\mathrm{sen}^2(\sigma)) = 0.$$

De aquí obtenemos que o bien r es 0, lo cual no nos interesa, o bien $r^2 - 4\cos^2(\sigma)\,\mathrm{sen}^2(\sigma) = (r + 2\cos(\sigma)\,\mathrm{sen}(\sigma))(r - 2\cos(\sigma)\,\mathrm{sen}(\sigma)) = 0$.
Por tanto tenemos dos expresiones para r, para cualquiera de las dos obtenemos la siguiente gráfica:

$$r = \mathrm{sen}(2\sigma)$$
$$r = -\mathrm{sen}(2\sigma)$$



2.1. Parametrización de Variedades Afines.

Nuestro objetivo es calcular los puntos de una variedad afín. Cuando tenemos un sistema de ecuaciones polinómicas de la forma $f_1 = \cdots = f_s = 0$, podemos tener un número de soluciones finito o infinito. En caso de que el número de soluciones sea finito, sólo tenemos que listar las soluciones del sistema y ya habríamos acabado. Pero, ¿qué pasa si hay infinitas soluciones?

Nuestro objetivo sería intentar expresar la forma de estos puntos, que, como veremos, se reduce a dar una parametrización de la variedad afín. Vamos a realizar un ejemplo en el que veremos cómo resolver un sistema polinómico sencillo. Además en este ejemplo vamos a parametrizar la solución.

Ejemplo 2.7. *Sea el anillo de polinomios $\mathbb{R}[x, y, z]$. Resolveremos el siguiente sistema polinómico:*

$$\left.\begin{array}{rcl} x+y+z & = & 1 \\ x+2y-z & = & 3. \end{array}\right\}$$

Geométricamente esto se corresponde con la intersección de los planos $\pi_1 : x + y + z = 1$ y $\pi_2 : x + 2y - z = 3$ cuyo corte da lugar a una recta. Es decir, es un sistema con infinitas soluciones.

Para resolverlo, aplicando reducción gaussiana a las filas obtenemos las siguientes ecuaciones equivalentes:

$$\left.\begin{array}{rcl} x+3z & = & -1 \\ y-2z & = & 2. \end{array}\right\}$$

Denotando $z = t$, con $t \in \mathbb{R}$, obtenemos la siguiente parametrización de las soluciones.

$$\left.\begin{array}{rcl} x & = & -1-3t \\ y & = & 2+2t \\ z & = & t. \end{array}\right\}$$

Hemos podido ver que cuando tenemos un sistema polinómico lineal, para parametrizar la variedad, tan solo tenemos que usar reducción gaussiana. Pero este trabajo se centrará en los sistemas polinómicos no lineales, que no son tan sencillos de resolver.

Llegados a este punto, las preguntas que nos surgen de manera natural son:

(i) Dado una superficie en forma paramétrica, ¿podemos conseguir su ecuación de forma implícita?

(ii) Dada una superficie en forma implícita, ¿podemos parametrizarla?

La respuesta a la segunda pregunta la hemos abordado brevemente en el ejemplo anterior en el que parametrizábamos la superficie con reducción gaussiana. Vamos a ver a continuación cómo funciona la implicitación, esto es: A partir de la parametrización de una superficie, intentar conseguir sus ecuaciones implícitas. La idea intuitiva sería eliminar los parámetros y que la ecuación final sólo dependa de las variables. Vamos a ver cómo funciona en un ejemplo:

Ejemplo 2.8. *Dada la parametrización de la siguiente curva, vamos a calcular su ecuación implícita:*

$$\left.\begin{array}{rcl} x & = & 1+t \\ y & = & 1+t^2. \end{array}\right\}$$

Usando la primera ecuación obtenemos $t = x - 1$. Sustituyendo en la segunda ecuación vemos que $y = x^2 - 2x + 2$. Por otro lado si decimos que $x = 1 + t$ y lo sustituimos en $y = x^2 - 2x + 2$ llegamos a que $y = 1 + t^2$, con lo que la curva anterior es la variedad $y = x^2 - 2x + 2$.

Hemos introducido lo que es una variedad afín y hemos visto ejemplos de la interpretación geométrica de las soluciones de ecuaciones polinómicas. En esta sección vamos a ver qué es un ideal así como algunas propiedades y resultados y resultados que nos serán de utilidad en el trabajo. Los ideales son muy importantes, ya que nos van a dar el lenguaje y las herramientas apropiadas para poder manejar y estudiar las variedades. Además aprenderemos más adelante a calcular las soluciones de los sistemas de polinomios haciendo uso de ellos, por ello debemos conocerlos y aprender a manipularlos. También recordaremos el algoritmo de la división para polinomios en una variable y el cálculo del máximo común divisor.

Por último, una vez definido todo, realizaremos ejemplos para la mejor compresión del tema.

2.2. Ideales.

Definición 2.9. *Un subconjunto $I \subseteq K[x_1, ..., x_n]$ es un* ideal *si cumple:*

(i) *Dado $f \in I$, entonces $-f \in I$ (cerrado por Opuestos)*

(ii) *Dados $f, g \in I$, entonces $f + g \in I$ (cerrado por la suma)*

(iii) *Dados $f \in I$ y $h \in K[x_1, ..., x_n]$ entonces $hf \in I$ (cerrado por producto de elementos del anillo)*

Lema 2.10. *Generadores y Base de un Ideal*

(i) *Sean $f_1, ..., f_s \in K[x_1, ..., x_n]$, entonces:*

$$\langle f_1, ..., f_s \rangle = \{h_1 f_1 + \cdots + h_s f_s : \ h_1, \cdots, h_s \in K[x_1, \cdots, x_n]\}$$

es un ideal en $K[x_1, ..., x_n]$. Nosotros los llamaremos, ideal generado por $f_1, ..., f_s$.

(ii) *Si existen $f_1, ..., f_s \in K[x_1, ..., x_n]$ tal que $I = \langle f_1, ..., f_s \rangle$ diremos que el ideal está finitamente generado, y llamaremos a $\{f_1, ..., f_s\}$ base del ideal.*

Al igual que vimos la interpretación geométrica de las soluciones, vamos a ver la relación de estos ideales con nuestros sistemas de ecuaciones.

Sea el ideal$\langle f_1, ..., f_s \rangle$ con $f_1, ..., f_s \in K[x_1, ..., x_n]$ y el sistema de ecuaciones:

$$\left.\begin{array}{rcl} f_1 & = & 0 \\ f_2 & = & 0 \\ \vdots & & \\ f_s & = & 0. \end{array}\right\}$$

Cuando resolvemos sistemas de ecuaciones lineales obtenemos unos equivalentes multiplicando las ecuaciones por escalares y sumándolas. Como estamos en el anillo $K[x_1, ..., x_n]$ las multiplicaremos por elementos del anillo. Por tanto, nuestras nuevas ecuaciones tendrán como poco las mismas soluciones que las originales. Así, si multiplico la primera ecuación por $h_1 \in K[x_1, ..., x_n]$, la segunda por $h_2 \in K[x_1, ..., x_n]$... y así sucesivamente, obtenemos una nueva ecuación:

$$h_1 f_1 + h_2 f_2 + ... + h_s f_s = 0$$

que es una *consecuencia polinómica* de nuestro sistema.

Definición 2.11. *Consecuencia Polinómica*
Sea el ideal $\langle f_1, ..., f_s \rangle$ *con* f_i, h_i *en* $K[x_1, ..., x_n]$. *Llamamos* consecuencias polinómicas *a los polinomios:*

$$h_1 f_1 + h_2 f_2 + ... + h_s f_s.$$

Ejemplo 2.12. *En el ejemplo 2.8 vimos cómo a partir de la parametrización:*

$$\left.\begin{array}{rcl} x & = & 1 + t \\ y & = & 1 + t^2, \end{array}\right\}$$

obteniamos la curva $y = x^2 - 2x + 2$, *Ahora vamos a volver a hacer el ejemplo usando las ideas anteriores de ideales.*

Primero, escribimos las ecuaciones de la siguiente forma:

$$\left.\begin{array}{rcl} x - 1 - t & = & 0 \\ y - 1 - t^2 & = & 0. \end{array}\right\}$$

Para cancelar la t *en ambas ecuaciones multiplicamos la primera por* $x - 1 + t$ *y a la segunda por* -1, *sumándolas obtenemos la ecuación:*

$$x^2 - 2x + 2 - y = 0,$$

que es la misma ecuación a la que habíamos llegado anteriormente.

Notemos que $x^2 - 2x + 2 - y \in \langle x - 1 - t, y - 1 - t^2 \rangle$, *y recordemos que este último es el conjunto de todas las posibles "consecuencias polinómicas"de los generadores del ideal.*

El siguiente resultado nos da la idea clave para resolver sistemas de ecuaciones polinómicos: encontrar otros polinomios que generan el mismo ideal pero cuyas ecuaciones son mas fáciles de resolver.

Proposición 2.13. *Distintas bases de un mismo ideal generan una misma variedad*
Si $\{f_1, ..., f_s\}$ *y* $\{g_1, ...g_t\}$ *son bases de un mismo ideal* I *en* $K[x_1, ..., x_n]$, *es decir, que* $I = \langle f_1, ...f_s \rangle = \langle g_1, ...g_t \rangle$, *entonces* $\mathbb{V}(f_1, ..., f_s) = \mathbb{V}(g_1, ..., g_t)$

Demostración. Probaremos que $\mathbb{V}(f_1, ..., f_s) = \mathbb{V}(g_1, ..., g_t)$ por doble contenido. Aunque en realidad sólo haremos el contenido de izquierda a derecha por que el otro se haace igual. Los puntos de $\mathbb{V}(f_1, ..., f_s)$ cumplen que son raíces de

cada f_i con $i = 0, 1, ..., s$. Como $\{f_1, ..., f_s\}$ es una base del ideal I entonces cada polinomio g_k puede ser escrito de la forma

$$g_k = \sum_{i=0}^{i=s} h_i f_i$$

con h_i un polinomio de $K[x_1, ..., x_n]$. Por tanto los puntos de $\mathbb{V}(f_1, ..., f_s)$ también serán raíces de g_k para cada $K = 1, ..., t$.

□

Definición 2.14. *Sea $\mathbb{V} \subseteq K^n$ una variedad afín. Entonces el conjunto:*

$$I(\mathbb{V}) = \{f \in K[x_1, ..., x_n] : f(a_1, ..., a_n) = 0,\ \forall (a_1, ..., a_n) \in \mathbb{V}\}$$

Lema 2.15. *Si $\mathbb{V} \subseteq K^n$ es una variedad afín, entonces $I(\mathbb{V})$ es un ideal, al que llamamos ideal de $\mathbb{V}$*

Lema 2.16. *Si $f_1, ..., f_s \in K[x_1, ..., x_n]$, entonces $\langle f_1, ..., f_s \rangle \subseteq I(\mathbb{V}(f_1, ..., f_s))$. La igualdad no siempre ocurre.*

Proposición 2.17. *Sea $\mathbb{V}$ y $\mathbb{W}$ variedades afines en K^n. Entonces:*

(i) $V \subseteq W \Leftrightarrow I(W) \subseteq I(V)$

(ii) $V = W \Leftrightarrow I(V) = I(W)$

Ahora bien, ¿cómo podemos ver que un ideal está contenido en otro?. Veamos el siguiente resultado.

Proposición 2.18. *Sea $I \subseteq K[x_1, ..., x_n]$ un ideal, y $f_1, ..., f_s \in K[x_1, ..., x_n]$. Entonces:*

$$f_1, ..., f_s \in I \Leftrightarrow \langle f_1, ..., f_s \rangle \subseteq I$$

2.3. Polinomios en una variable.

En este apartado daremos un rápido repaso a las nociones básicas de polinomios en una variable, enlazando esto con lo visto anteriormente sobre ideales. Veremos la definición formal de *máximo común divisor* de varios polinomios y un método para calcularlo. Esto también lo haremos cuándo un polinomio está en un ideal.

Estos conceptos nos serán muy útiles más adelante, ya que todos los cálculos estarán basados en estas operaciones.

Comenzamos viendo algunos resultados interesantes.

Proposición 2.19. *El Algoritmo de la División*
Sea K un cuerpo y $g \in K[x]$ no nulo. Entonces, todo polinomio $f \in K[x]$ puede ser escrito de la siguiente forma:

$$f = qg + r,$$

donde $q, r \in K[x]$ y $\partial(r) < \partial(g)$, ($\partial(r)$ denota el grado del polinomio). Además q,r son únicos y existe un algoritmo para calcularlos.

Proposición 2.20. *El Algoritmo de la División*
Input: g,f
Output: q,r
q:=0; r:= f

WHILE $r \neq 0$ *AND LT(g) divida a LT(r) DO*
$q := q + \frac{LT(r)}{LT(g)}$
$r := r - \frac{LT(r)}{LT(g)} g$

Corolario 2.21. *Si K es un cuerpo y* $f \in K[x]$*, un polinomio no nulo. Entonces f tiene a lo más* $\delta(f)$ *soluciones en el cuerpo K*

Corolario 2.22. *Si K es un cuerpo, entonces todo ideal en* $K[x]$ *puede ser escrito de la forma* $\langle f \rangle$*, para algún* $f \in K[x]$*. Además f es único salvo unidades de K[x], es decir, salvo escalares no nulos de K.*

Demostración. Sea un ideal $I \subseteq K[x]$. Si $I = \{0\}$ entonces tomando $f = 0$ habríamos acabado. Por otro lado, sea f un polinomio contenido en I no nulo cuyo grado es el más pequeño. Veamos que $I = \langle f \rangle$. La inclusión de derecha a izquierda es obvia ya que $f \in I$. Vamos a demostrar la otra inclusión. Sea un $g \in I$. Por el algoritmo de la división tenemos que $g = qf + r$ donde o bien $r = 0$ o bien el grado de r es más pequeño que el grado de f. Ya que I es un ideal entonces $qf \in I$ y así $g - qf = r \in I$, lo cual contradice que f sea de grado más pequeño de todos los elementos de I. Por lo tanto $r = 0$, entonces $g = qf \in \langle f \rangle$. Esto prueba que $I = \langle f \rangle$. Para ver la unicidad supongamos que $\langle f \rangle = \langle g \rangle$. Entonces $f \in \langle g \rangle$ implica que $f = hg$ para algún polinomio h. Así

$$\delta(f) = \delta(h) + \delta(g),$$

por lo que $\delta(f) \geq \delta(g)$. El mismo argumento intercambiando los papeles de g y f nos lleva a que $\delta(g) \geq \delta(f)$, por lo tanto $\delta(f) = \delta(g)$. Esto implica que h sea una constante no nula.

□

Este resultado nos dice que para resolver sistemas polinómicos en una variable determinados por un conjunto de polinomios, basta resolver el polinomio que general el ideal que determina el conjunto de polinomios. Este polinomio que genera el ideal es el máximo común divisor de los polinomios como veremos más adelante.

Definición 2.23. *Máximo común divisor*

El máximo común divisor *de los polinomios* $f_1, ..., f_s \in K[x]$ *es un polinomio* $h \in K[x]$ *tal que:*

(i) *h divide a* $f_1, ..., f_s$ *(común divisor).*

(ii) *Si* $\exists p \in K[x]$ *tal que p divide a* $f_1, ..., f_s$ *entonces p divide a h (máximo).*
Al polinomio h que cumple estas propiedades lo denotamos. $MCD(f_1, ..., f_s)$

Proposición 2.24. *Sean* $f_1, ..., f_s \in K[x]$*, con* $s \geq 2$*. Entonces:*

(i) $MCD(f_1, ..., f_s)$ *existe y es único salvo unidades..*

(ii) *$MCD(f_1, ..., f_s)$ es un generador de $\langle f_1, ..., f_s \rangle$.*

(iii) *Si $s \geq 3$ entonces $MCD(f_1, ..., f_s) = MCD(f_1, MCD(f_2, ..., f_s))$.*

(iv) *Existe un algoritmo para encontrar el $MCD(f_1, ..., f_s)$.*

Corolario 2.25. *Sea K un cuerpo y $f_1, ..., f_s, g \in K[x]$. Entonces $g \in \langle f_1, ..., f_s \rangle$ si y solo si el $resto(g, MCD(f_1, ..., f_s)) = 0$*

Es decir, que en una variable el polinomio que genera el ideal que determinan varios polinomios es el MCD. Y un polinomio está en ese ideal si lo divide el MCD. Luego resolver un sistema de ecuaciones polinómicas determinado por polinomios en una variable equivale a resolver la ecuación que determina el MCD. Es cierto que con esto no ha terminado el problema si el MCD tiene grado mayor o igual que 5 (no hay fórmula para calcular las raíces, por la Teoría de Galois).

Para calcular el MCD de varios polinomios en una variable usaremos el *Algoritmo de Euclides*.

Su funcionamiento es simple. Tan solo tendremos que hacer sucesivas divisiones de polinomios hasta llegar a una división exacta. A continuación mostramos el pseudocódigo del algoritmo.

Proposición 2.26. *Algoritmo de euclides para MCD*
Input: f,g
Output: h
h:=f; s:= g

WHILE $s \neq 0$ DO
$resto := resto(h, s)$
$h := s$
$s := resto$

Vamos a hacer dos ejemplos en los que calcularemos el MCD de polinomios en una variable. Primero lo realizaremos con dos polinomios y luego con tres polinomios.

Ejemplo 2.27. *Vamos a calcular el MCD de los polinomios $f = x^4 - 1$ y $g = x^6 - 1$*
Comenzamos dividiendo lo polinomios:

$A_1 = x^2$

$x^4 - 1$ | $x^6 - 1$
$x^6 - x^2$
$x^2 - 1$

Ahora, como dice el algoritmo de euclides para el MCD, divido el dividendo entre el resto:

$A_1 = x^2 + 1$

$x^2 - 1$	$x^4 - 1$
	$x^4 - x^2$
	$x^2 - 1$
	$x^2 - 1$
	0

Como el resto es cero, el $MCD(x^6 - 1, x^4 - 1)$ *es el último resto no nulo, es decir* $MCD(x^6 - 1, x^4 - 1) = x^2 - 1$.

Ejemplo 2.28. *Calcularemos el MCD de los polinomios* $f = x^4 - 1, g = x^6 - 1, h = x^2 + 1$ *Voy a usar el Algoritmo de Euclides. Los ordeno por su grado (notar que no es necesario ordenarlos por su grado, lo hacemos por comodidad).*

Así tenemos:

$$MCD(g, f, h) = MCD(x^6 - 1, x^4 - 1, x^2 - 1) =$$
$$MCD(x^6 - 1, MCD(x^4 - 1, x^2 - 1)) = MCD(x^6 - 1, x^2 - 1)$$

Para calcular este último MCD usamos el algoritmo:

$A_1 = x^4 + x^2 + 1$

$x^2 - 1$	$x^6 - 1$
	$x^6 - x^4$
	$x^4 - 1$
	$x^4 - x^2$
	$x^2 - 1$
	$x^2 - 1$
	0

Como el resto es cero significa que $x^2 - 1$ *divide a* $x^6 - 1$ *y por tanto* $MCD(f, g, h) = x^2 - 1$

Ejemplo 2.29. *Calcularemos el MCD de los polinomios* $f = x^3 - 3x + 2, g = x^6 - 1, h = x^4 - 1$ *Para este ejemplo, vamos a proceder de la misma forma que en el ejemplo anterior, con la salvedad de que no ordenaremos los polinomios por su grado. Aprovecharemos las cuentas anteriores.*

$$MCD(f, g, h) = MCD(f, MCD(g, h)) = MCD(f, x^2 - 1)$$

$A_1 = x$

$x^2 - 1$	$x^3 - 3x + 2$
	$x^3 - x$
	$-2x + 2$

$A_1 = \frac{-x}{2} - \frac{1}{2}$

-2x + 2	$x^2 - 1$
	$x^2 - x$
	x - 1
	x - 1
	0

Por tanto tenemos que $MCD(f, g, h) = x - 1$

A continuación veremos si el polinomio $x^3 + 4x^2 + 3x - 7$ *está en el ideal* $\langle f, g, h\rangle$*, es decir si*

$$x^3 + 4x^2 + 3x - 7 \in \langle x^3 - 3x + 2, x^6 - 1, x^4 - 1\rangle.$$

Para hacer este ejemplo, nos basta con usar el Corolario 2.25.

$$\langle x^3 - 3x + 2, x^6 - 1, x^4 - 1\rangle = \langle MCD(x^3 - 3x + 2, x^6 - 1, x^4 - 1)\rangle = \langle x - 1\rangle.$$

Procedemos con la división:

$A_1 = x^2 + 5x + 8$

x-1	$x^3 + 4x^2 + 3x - 7$
	$x^3 - x^2$
	$5x^2 + 3x - 7$
	$5x^2 - 5x$
	$8x - 7$
	$8x - 8$
	1

Como la división no es exacta, entonces podemos afirmar que

$$x^3 + 4x^2 + 3x - 7 \notin \langle x^3 - 3x + 2, x^6 - 1, x^4 - 1\rangle.$$

3. Bases de Groebner.

En la sección 2 hemos introducido la geometría de las variedades afines y el álgebra de los anillos de polinomios en n variables, aunque nos hemos centrado en anillos de polinomios en una sola variable.

El método de las bases de Groebner nos permitirá resolver sistemas polinómicos. Hasta ahora, calculábamos el generador de un ideal del anillo de polinomios en una variable mediante el MCD. Gracias a las bases de Groebner conseguiremos un generador o base de generadores de los ideales de polinomios en varias variables, y con ellos trataremos de resolver sistemas de ecuaciones polinómicas.

Uno de los primeros problemas que abordaremos es el orden de los monomios. Cuando resolvemos un sistema de ecuaciones polinómicas en una variable, usamos el MCD para resolverlo y con ello la división. Pero para comenzar la división, necesitamos ordenar los monomios.

Por ejemplo, tenemos $f = xy^2 + x^2y \in \mathbb{R}[x, y]$ y queremos ordenar sus términos. Ambos monomios tienen grado 3, ¿qué criterio usamos?, ¿cómo los ordenamos? En el siguiente apartado resolveremos este problema.

3.1. Orden monomial y algoritmo de la división en n variables.

En este apartado estudiaremos con detenimiento las distintas formas de ordenar monomios. Cuando dividimos polinomios, comenzamos ordenando los monomios del dividendo y divisor por su grado. Ahora bien ¿cómo sabemos que monomio es el que más grado tiene?. La respuesta a esta pregunta viene dada por la siguiente definición.

Definición 3.1. *Orden monomial*
Un orden monomial $>$ *en el conjunto de monomios de* $K[x_1, ..., x_n]$ *es una relación de orden en el conjunto de monomios* x^α *donde* $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ *con* $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$*, o equivalentemente, cualquier relación de orden en* $\alpha \in \mathbb{N}^n$ *que cumple:*

(i) $>$ *es un orden total en* $\mathbb{N}^n$*.*

(ii) *Si* $\alpha > \beta$ *y* $\gamma \in \mathbb{N}^n$ *entonces* $\alpha + \gamma > \beta + \gamma$*.*

(iii) $>$ *es* un buen orden *en* $\mathbb{N}^n$*, es decir, que todo conjunto no vacío de* $\mathbb{N}^n$ *tiene un elemento que es el más pequeño bajo la relación* $>$*.*

Lema 3.2. *Una relación de orden* $>$ *en* $\mathbb{N}^n$ *es de buen orden si, y solo si, se cumple que cualquier sucesión estrictamente decreciente:*

$$\alpha(1) > \alpha(2) > \alpha(3) > \cdots$$

se termina en algún momento.

Corolario 3.3. *Sea* $>$ *una relación en* $\mathbb{N}^n$ *que cumple:*

(i) $>$ *Es un orden total en* $\mathbb{N}^n$*.*

(ii) *Si* $\alpha > \beta$ *y* $\gamma \in \mathbb{N}^n$ *entonces* $\alpha + \beta > \beta + \gamma$*.*

Entonces $>$ *es un buen orden si, y solo si,* $\alpha \geq 0, \forall \alpha \in \mathbb{N}^n$*.*

Definición 3.4. *Conceptos sobre un polinomio*
Sea $f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha$ *un polinomio distinto de cero en* $K[x_1, ..., x_n]$ *y sea* $>$ *un orden monomial.*

(i) *El* multigrado *de* f *respecto a* $>$ *es*

$$multigr(f) = max\{\alpha \in \mathbb{N}^n : \ a_\alpha \neq 0\}.$$

(ii) *El* coeficiente líder *de* f *es*

$$LC(f) = a_{multigr(f)} \in K.$$

(iii) *El* monomio líder *de* f *es*

$$LM(f) = x^{multigr(f)}.$$

(iv) *El* término líder *de* f *es*

$$LT(f) = LC(f) \cdot LM(f).$$

Una vez definidos estos conceptos, vamos a estudiar tres tipos de ordenes distintos para ordenar nuestros monomios. Existen muchos órdenes, incluso podríamos inventar orden, pero en este trabajo solo usaremos los más comunes.

Definición 3.5. *Orden Lexicográfico*
Sea $\alpha = (\alpha_1, ..., \alpha_n)$ *y* $\beta = (\beta_1, ..., \beta_n)$*. Diremos que* $\alpha >_{Lex} \beta$ *si, en el vector diferencia* $\alpha - \beta \in \mathbb{N}^n$*, el elemento más a la izquierda distinto de cero es positivo. En este caso escribiremos* $x^\alpha >_{Lex} x^\beta$ *o* $\alpha >_{Lex} \beta$*.*

Ejemplo 3.6. *Usamos el orden Lexicográfico para ordenar* $\alpha = (1, 2, 0), \beta = (0, 3, 4)$.
El vector resta $\alpha - \beta = (1, -1, -4)$ *Por tanto* $\alpha >_{Lex} \beta$*.*

Ejemplo 3.7. *Usaremos el orden Lexicográfico para ordenar* $\alpha = (3, 2, 4), \beta = (3, 2, 6)$.
Para ello, en vez de utilizar el vector resta, vamos a fijarnos en las componentes de los dos vectores. Comparando sus componentes una a una y en la primera componente más grande, el vector al que pertenezca esa componente, ese será el vector más grande. Las dos primeras componentes coinciden así que nos fijamos en la tercera componente, la tercera componente es más grande en el vector β*, por tanto:* $\beta >_{Lex} \alpha$

Notemos que en este orden, y en cualquiera de los que siguen partimos de un ordenamiento de las variables $x_1, ..., x_n$.

Definición 3.8. *Orden Lexicográfico Graduado*
Sea $\alpha, \beta \in \mathbb{N}^n$*. Diremos que* $\alpha >_{grlex} \beta$ *si:*

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i$$

En tal caso, diremos que $x^\alpha >_{grlex} x^\beta$ *o* $\alpha >_{grlex} \beta$*. Si ambos grados coinciden, romperemos el empate usando el orden Lexicográfico.*

Ejemplo 3.9. *Usaremos el orden Lexicográfico Graduado para ordenar* $\alpha = (1,4,3), \beta = (3,2,0)$.
Como $|\alpha| = 8 > |\beta| = 6$ *concluimos que* $\alpha >_{grLex} \beta$.

Ejemplo 3.10. *Usaremos el orden Lexicográfico Graduado para ordenar* $\alpha = (1,2,4), \beta = (1,1,5)$.
Podemos ver que $|\alpha| = |\beta| = 7$*, así que desempatamos usando el orden Lexicográfico. Nos fijamos en la segunda componente de* α *que es más grande que la segunda componente de* β *por tanto:* $\alpha >_{grlex} \beta$.

Definición 3.11. *Orden Lexicográfico Graduado Inverso*
Sean $\alpha, \beta \in \mathbb{N}^n$. *Diremos que* $\alpha >_{grevlex} \beta$ *si:*

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i \text{ o si}$$

$|\alpha| = |\beta|$ *y la entrada no nula mas a la izquierda del vector* $\alpha - \beta$ *es negativa.*

En tal caso diremos que $x^\alpha >_{grevlex} x^\beta$ *o* $\alpha >_{lex} \beta$.

Los órdenes grevlex y grLex se comportan de forma similar. Sólo varían en la forma de decidir cual es de mayor grado cuando el grado total de los vectores coinciden. Vamos a ver cómo funciona este último orden en el caso de empate.

Ejemplo 3.12. *Usaremos el orden Lexicográfico Graduado Inverso para ordenar* $\alpha = (1,2,4), \beta = (1,1,5)$.
Podemos ver que $|\alpha| = |\beta| = 7$*, así que desempatamos usando el orden Lexicográfico Graduado Inverso. Nos fijamos en la segunda componente de* β *que es más pequeña que la segunda componente de* α *por tanto:* $\beta >_{grLex} \alpha$.

La idea es que cuando usamos el grevlex para desempatar, no buscamos la componente más grande, si no la más pequeña.

3.2. El algoritmo de la división en varias variables.

Anteriormente vimos el algoritmo de la división para polinomios en una variable. En esta sección veremos cómo es la división de polinomios para n variables. Comenzaremos viendo dos ejemplos de esta división y, finalmente, daremos el algoritmo de la división para n variables. En general nosotros tendremos que dividir $f \in K[x_1, ..., x_n]$ por $f_1, ..., f_s \in K[x_1, ..., x_n]$. Como veremos esto significa expresar f de la forma

$$f = a_1 f_1 + ... + a_s f_s + r$$

donde los coeficientes $a_1, ..., a_s$ y el resto r están en $K[x_1, ..., x_n]$. Los órdenes que definimos anteriormente serán clave a la hora de caracterizar el resto.

La idea básica del algoritmo es la misma que en una variable: queremos cancelar el término líder de f (respecto al orden monomial fijado) multiplicando el divisor f_i por otro polinomio y restándolos.

Ejemplo 3.13. *Vamos a dividir* $f = xy^2+1$ *por* $f_1 = xy+1$ *y* $f_2 = y+1$ *usando el orden Lexicográfico con* $x > y$*. A diferencia de la división en una variable, ahora vamos a dividir entre varios divisores. Usaremos la siguiente notación*

	$A_1=$
	$A_2=$
$xy+1$	xy^2+1
$y+1$	

Los términos líder $LT(f_1) = xy$ y $LT(f_2) = y$ dividen a $LT(f) = x^2y$. Así pues, podemos empezar con el primero para dividir:

	$A_1= y$
	$A_2=$
$xy+1$	xy^2+1
$y+1$	xy^2+y
	$-y+1$

Si repetimos el proceso en -y+1, ahora debemos usar f_2 ya que $LT(f_1) = xy$ no divide a $LT(-y+1) = -y$. Así obtenemos:

	$A_1= y$
	$A_2= -1$
$xy+1$	xy^2+1
$y+1$	xy^2+y
	$-y+1$
	$-y-1$
	2

Como ningún término líder divide a 2, hemos acabado. De esta forma obtenemos

$$xy^2 + 1 = y(xy + 1) + (-1)(y + 1) + 2$$

Ejemplo 3.14. *En este ejemplo, encontraremos un problema que no ocurre en una variable. Vamos a dividir $f = x^2y + xy^2 + y^2$ por $f_1 = xy - 1$ y $f_2 = y^2 - 1$. Comenzamos con el mismo procedimiento que hemos usado en el ejemplo anterior, obteniendo*

A_1=x +y

xy -1	$x^2y + xy^2 - y^2$
y^2 -1	x^2y - x
	$xy^2 + x - y^2$
	xy^2 - y
	x - y^2 + y

Notemos que ni $LT(f_1) = xy$ ni $LT(f_2) = y^2$ divide a $LT(x + y^2 + y) = x$. Pero $x+y^2+y$ no es el resto ya que $LT(f_2)$ divide a y^2. Así, si nosotros movemos x al resto, podemos seguir dividiendo. Para implementar la idea creamos una columna r, a la derecha de la división, donde iremos colocando los elementos del resto.

A_1=x +y

A_2=1

		r
xy -1	$x^2y + xy^2 - y^2$	
y^2 -1	x^2y - x	
	$xy^2 + x - y^2$	
	xy^2 - y	
	x - y^2 + y ⟶	x
	y^2 + y	

Ahora podemos seguir dividiendo. Si podemos dividir por $LT(f_1)$ o $LT(f_2)$ continuaremos de la forma usual, y si no podemos moveremos el término líder a la columna del resto.

A_1=x +y

A_2=1

		r
xy -1	$x^2y + xy^2 - y^2$	
y^2 -1	x^2y - x	
	$xy^2 + x - y^2$	
	xy^2 - y	
	x - y^2 + y ⟶	x
	y^2 + y	
	$y^2 - 1$	
	y + 1 ⟶	x + y
	1 ⟶	x + y + 1
	0	

Así obtenemos

$$x^2y + xy^2 + y^2 = (x + y)(xy - 1) + 1(y^2 - 1) + x + y + 1.$$

Observemos que el resto es suma de monomios que no son divisibles ni por $LT(f_1)$ ni por $LT(f_2)$.

Hemos ejemplarizado cómo funciona el algoritmo de la división. Vamos a dar a continuación la forma general de dicho algoritmo.

Teorema 3.15. *Algoritmo de la División*
Fijado un orden $>$ en $\mathbb{N}^n$, y dada $F = (f_1, ..., f_s)$ una serie de polinomios en $K[x_1, ..., x_n]$, entonces todo polinomio f en $K[x_1, ..., x_n]$ puede ser escrito de la siguiente forma

$$f = a_1 f_1 + ... + a_s f_s + r,$$

donde $a_1, r \in K[x_1, ..., x_n]$ y o bien $r=0$ o r es una suma de monomios los cuales ninguno es divisible por ninguno de los monomios $LT(f_1), ..., LT(f_s)$. Denotamos por r al resto de la división de f por F. Además si $a_i f_i \neq 0$ entonces

$$multigr(f) \leq multigr(a_i f_i).$$

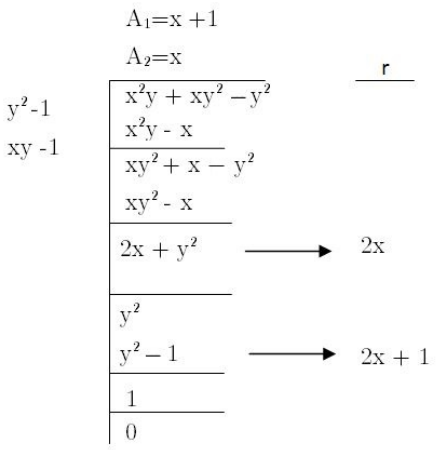
Una generalización del teorema puede ser dada por el siguiente código:

Input: $f_1, ..., f_s, f$
Output: $a_1, ..., a_s, r$
a_1:=0;...; a_s:=0 r:= 0
p:= f

WHILE $p \neq 0$ DO
i:=1
división := false
WHILE $i \leq s$ AND división=false DO
IF $LT(f_i)$ divide a (p) THEN
$a_i := a_i + LT(p)/LT(f_i)$
$p := p - (LT(p)/LT(f_i))f_i$
división:= true
ELSE
i:= i+1
IF división = false THEN
r:= r + LT(p)
p:= p - LT(p)

Concluimos esta sección preguntándonos si el algoritmo de la división en n variables tiene las mismas propiedades que el de una variable. Desafortunadamente no es así. La propiedad más importante de la división en una variables es la unicidad del resto. Para ver cómo puede fallar esta propiedad en n variables, vamos a estudiar el siguiente ejemplo.

Ejemplo 3.16. *Vamos a dividir $F = x^2y+xy^2+y^2$ por $f_1 = y^2-1$ y $f_2 = xy-1$ usando el orden Lexicográfico con $x > y$. Estos polinomios son los mismos que en el Ejemplo 3.14., con la excepción de que vamos a cambiar el orden de los divisores. Aplicando el algoritmo de la división en n variables obtenemos*



Así

$$x^2y + xy^2 + y^2 = (x+1)(y^2-1) + x(xy-1) + 2x + 1.$$

Si comparamos el resto obtenido en este ejemplo, $2x+1$, con el resto del Ejemplo 3.14, $x+y+1$, observamos que son distintos. Esto demuestra que ni el resto ni los a_i son únicos. Si alteramos el orden de los divisores, estos podrían cambiar. A lo largo del trabajo comprobaremos que cuando dividimos entre los generadores de una base de Groebner, el resto sí es único, independientemente del orden de los divisores. Para definir estas bases apropiadamente necesitamos los conceptos y resultados de las siguientes subsecciones.

3.3. Lema de Dickson.

Definición 3.17. *Ideal monomial*
Un ideal $I \subseteq K[x_1, ..., x_n]$ es un ideal monomial si existe $A \subseteq \mathbb{N}^n$ tal que I consiste en todos los polinomios formados por sumas finitas de la forma $\sum_{\alpha \in A} h_\alpha x^\alpha$, donde $h_\alpha \in K[x_1, ..., x_n]$. En este caso denotamos al ideal de monomios por $I = \langle x^\alpha | \alpha \in A \rangle$. A cada elemento x^α con $\alpha \in A$, lo llamaremos generador del ideal de monomios.

Lema 3.18. *Sea $I = \langle x^\alpha | \alpha \in A \rangle$ un ideal monomial y x^β un monomio. Entonces son equivalentes:*

$$x^\beta \in I \text{ si y solo si } x^\beta \text{ es divisible por algún } x^\alpha \in I$$

Lema 3.19. *Sea I un ideal monomial y $f \in K[x_1, ..., x_2]$. Entonces es equivalente:*

(i) $f \in I$.

(ii) *Cada término de $f \in I$.*

(iii) *f es una K-combinación lineal de monomios de I.*

Teorema 3.20. *Lema de Dickson*
Sea $I = \langle x^\alpha | \alpha \in A \rangle \subseteq K[x_1, ..., x_n]$ un ideal monomial. Entonces I puede ser escrito de la forma $I = \langle x^{\alpha(1)}, ..., x^{\alpha(s)} \rangle$ donde $\alpha(1), ..., \alpha(s) \in A$. En particular I tiene un número finito de generadores.

Demostración.

Realizaremos la demostración por inducción en n, el número de variables. Si n=1, entonces I es generado por los monomios x_1^α donde $\alpha \in A$. Sea $\beta \leq \alpha$ el elemento más pequeño de A. Entonces $\beta \leq \alpha$ para cualquier elemento α de A. Por esto, x_1^β divide a todos los demás generadores x_1^α así que tenemos que $I = \langle x_1^\beta \rangle$.

Ahora asumimos que $n > 1$ y que el teorema es cierto para $n-1$. Trabajaremos con las variables $x_1, ..., x_{n-1}, y$, así nuestros monomios estarán en $K[x_1, ..., x_{n-1}, y]$ y podemos escribirlos de la forma $x^\alpha y^m$ donde

$$\alpha = (\alpha_1, ..., \alpha_{n-1}) \in \mathbb{N}^{n-1} \text{ y } m \in \mathbb{N}.$$

Supongamos que $I \subseteq K[x_1, ..., x_{n-1}, y]$ es un ideal monomial. Para encontrar sus generadores, sea J el ideal generado por los monomios x^α los cuales $x^\alpha y^m \in I$ para algún $m \geq 0$. Ya que J es un ideal de monomios, nuestra hipótesis de inducción implica que un número finito de x^α generan J, es decir $J = \langle x^{\alpha(1)}, ..., x^{\alpha(s)} \rangle$.

Para cada i entre 1 y s, la definición de J nos dice que $x^{\alpha(1)} y^{m_i} \in I$ para algún $m_i \geq 0$. Sea m el más grande de esos m_i. Entonces para cada k entre 0 y m-1, consideramos el ideal $J_k \subseteq K[x_1, ..., x_{n-1}]$ generado por los monomios x^β tal que $x^\beta y^k \in I$. Podemos pensar que J_k es una parte de I generado por los monomios que tienen la variable y elevada a la k-ésima potencia. Usando nuestra hipótesis de inducción otra vez, J_k tiene un conjunto monomial finito de generadores es decir $J_k = \langle x^{\alpha_k(1)}, ..., x^{\alpha_k(s)} \rangle$.

Veamos que I es generado por los monomios de la siguiente lista:

$$\begin{aligned} &J : x^{\alpha(1)} y^m, ..., x^{\alpha(s)} y^m, \\ &J_0 : x^{\alpha_0(1)}, ..., x^{\alpha_0(s_0)}, \\ &J_1 : x^{\alpha_1(1)} y, ..., x^{\alpha_1(s_1)} y, \\ &\vdots \\ &J_{m-1} : x^{\alpha_{m-1}(1)} y^{m-1}, ..., x^{\alpha_{m-1}(s_{m-1})} y^{m-1}, \end{aligned}$$

Primero notemos que todos los monomios en I son divisibles por uno de la lista anterior. Para ver esto, sea $x^\alpha y^p \in I$. Si $p \geq m$, entonces $x^\alpha y^p$ es divisible por algún $x^{\alpha(i)} y^m$, debido a la construcción de J. Por otro lado, si $p \leq m-1$, entonces $x^\alpha y^p$ es divisible por algún $x^{\alpha_p(j)} y^p$ debido a la construcción de J_p. Debido al Lema 3.18. los monomios de antes generan un ideal que contiene los mismos monomios que I, por lo tanto son los mismos ideales.

Para completar la demostración del teorema, necesitamos que el conjunto finito de generadores pueda ser elegido del conjunto de generadores del ideal dado. Por el párrafo anterior sabemos que $I = \langle x^{\beta(1)}, ..., x^{\beta(s)} \rangle$ para algún monomio $x^{\beta(i)}$ en I. Ya que $x^{\beta(i)} \in I = \langle x^\alpha : \alpha \in A \rangle$, el lema 3.18. nos dice que cada $x^{\beta(i)}$ es divisible por $x^{\alpha(i)}$ para algún $\alpha(i) \in A$. De aquí tenemos tenemos que $I = \langle x^{\alpha(1)}, ..., x^{\alpha(s)} \rangle$. Esto completa la demostración.

□

3.4. Teorema de la Base de Hilbert.

Definición 3.21. *Sea $I \subseteq K[x_1, ..., x_n]$ un ideal distinto del vacío.*

(i) *Denotamos al conjunto de los términos líder de los elementos de I por LT(I). Y lo definimos como:*

$$LT(I) = \{cx^\alpha : \exists\ f \in I \text{ con } LT(f) = cx^\alpha\}.$$

(ii) *Denotamos al ideal generado por los elementos de LT(I) por* $\langle LT(I)\rangle$.

Ejemplo 3.22. *Sea* $I = \langle f_1, f_2\rangle$ *con* $f_1 = x^3 - 2xy$ *y* $f_2 = x^2y - 2y^2 + x$, *usamos el orden Lexicográfico Graduado para ordenar los polinomios* $f_1, f_2 \in K[x, y]$. *Podemos ver que:*

$$x \cdot f_2 - y \cdot f_1 = x^2$$

Por lo tanto $x^2 \in I$. *Así es evidente que* $x^2 = LT(x^2) \in \langle LT(I)\rangle$. *Sin embargo* x^2 *no es divisible por el* $LT(f_1) = x^3$ *ni por el* $LT(f_2) = x^2y$. *Por tanto aplicando el Lema 3.18 tenemos que* $x^2 \notin \langle LT(f_1), LT(f_2)\rangle$

Proposición 3.23. *Sea* $I \subseteq K[x_1, ..., x_n]$ *un ideal. Entonces:*

(i) $\langle LT(I)\rangle$ *es un ideal monomial.*

(ii) *Existen* $g_1, ..., g_t \in I$ *tales que* $\langle LT(I)\rangle = \langle LT(g_1), ..., LT(g_t)\rangle$. *Es decir* $\langle LT(I)\rangle$ *posee un número finito de generadores.*

Una vez hemos enunciado esta proposición, ya estamos en condiciones de enunciar uno de los Teoremas más importantes del trabajo, conocido como el Teorema de la base de Hilbert.

Figura 2: David Hilbert (1862–1943) fue uno de los matemáticos mas influyentes del siglo XIX y XX. Desarrolló los *espacios Hilbert*.

Teorema 3.24. *Teorema de la base de Hilbert*
Todo ideal monomial $I \subseteq K[x_1, ..., x_n]$ *tiene una familia finita de generadores, es decir,* $I = \langle g_1, ..., g_t\rangle$ *para algunos* $g_1, ..., g_t \in I$.

Demostración.

Si $I = \{0\}$, tomamos el conjunto de generadores $\{0\}$ y ya habríamos acabado. Por otro lado, si I contiene algún elemento distinto de 0, entonces un conjunto generador del ideal I puede ser construido de la siguiente forma. Por la Proposición 3.23. existen $g_1, ..., g_t \in I$ tales que $\langle LT(I)\rangle = \langle LT(g_1), ..., LT(g_t)\rangle$.

Veamos que $I = \langle g_1, ..., g_t\rangle$. Tenemos que $\langle g_1, ..., g_t\rangle \subseteq I$ ya que cada $g_i \in I$. Para el otro contenido, sea $f \in I$ cualquier elemento de I. Si aplicamos el algoritmo de la división para dividir f por $\{g_1, ..., g_t\}$ obtenemos la siguiente expresion

$$f = a_1g_1 + ... + a_tg_t + r$$

donde ningún término de r es divisible por $LT(g_1), ..., LT(g_t) \in I$. Si $r \neq 0$, entonces $LT(r) \in \langle LT(I)\rangle = \langle LT(g_1), ..., LT(g_t)\rangle$ y por el Lema 3.18. tenemos que LT(r) debe ser divisible por algún $LT(g_i)$. Esto contradice que r sea un resto y por tanto r debe ser 0. Así

$$f = a_1g_1 + ... + a_tg_t \in \langle g_1, ..., g_t\rangle,$$

lo cual demuestra que $I \subseteq \langle g_1, ..., g_t\rangle$. □

Figura 3: Bruno Buchberger (1942–actualidad) es un matemático austriaco que destacó en temas relacionados con el álgebra computacional. Sus principales contribuciones a la ciencia fueron las bases de Groebner.

3.5. Bases de Groebner.

Con el contenido de las dos subsecciones anteriores podemos ya definir el concepto de base de Groebner. Esas son las bases que nos interesan de los ideales generados por polinomios de varias variables.

Definición 3.25. *Base de Groebner*
Dado un orden monomial y un ideal $I \in K[x_1, ..., x_n]$*, un subconjunto finito* $G = \{g_1, ..., g_n\} \subseteq I$ *se llama* base de Groebner *si cumple:*

$$\langle LT(g_1), ..., LT(g_n)\rangle = \langle LT(I)\rangle$$

Notar que también podemos decir "Una familia $g_1, ..., g_n \subseteq I$ es una base de Groebner de I si y solo si el LT de cualquier elemento de I es divisible por algún $LT(g_i)$". Este hecho lo demostraremos en la Observaación 3.31. A continuación, vemos una consecuencia directa del Teorema de la base de Hilbert.

Corolario 3.26. *Fijado un orden monomial, entonces todo ideal* $I \subseteq K[x_1, ..., x_n]$ *distinto del vacío posee una base de Groebner. Además cualquier base de Groebner de un ideal I, es base de I.*

Teorema 3.27. *Sean* $I_1 \subseteq I_2 \subseteq ...$ *una sucesión creciente de ideales en* $K[x_1, ..., x_n]$*. Entonces existe* $N \geq 1$ *para el cual se cumple :*

$$I_N = I_{N+1} = I_{N+2} = ...$$

Definición 3.28. *Sea* $I \subseteq K[x_1, ..., x_n]$ *un ideal. Denotamos por* $\mathbb{V}(I)$ *al conjunto :*

$$\mathbb{V}(I) = \{(a_1, ..., a_n) \in K^n \ : \ f(a_1, ..., a_n) = 0 \ \forall f \in I\}.$$

Proposición 3.29. $\mathbb{V}(I)$ *es una variedad afín. En particular si* $I = \langle f_1, ..., f_s\rangle$*, entonces* $\mathbb{V}(I) = V(f_1, ..., f_s)$*.*

Para terminar este apartado vamos a ver unos cuantos ejemplos y observaciones.

Ejemplo 3.30. *Sea* $I = \langle g_1, g_2, g_3\rangle \subseteq \mathbb{R}[x, y, z]$ *donde* $g_1 = xy^2 - xz + y$*,* $g_2 = xy - z^2$ *y* $g_3 = x - yz^4$*. Usando el orden Lexicógrafico damos un ejemplo de* $g \in I$ *que cumple que* $LT(g) \notin \langle LT(g_1), LT(g_2), LT(g_3)\rangle$*.*

Debemos buscar un polinomio g tal que:

$$LT(g) \notin \langle LT(g_1), LT(g_2), LT(g_3)\rangle = \langle xy^2, xy, x\rangle.$$

Nuestra estrategia para encontrar dicho polinomio será buscar uno que solo tenga las variables y ó z. Sean los polinomios $h_1 = y$, $h_2 = z$, $h_1, h_2 \in mathbbR[x, y, z]$. El polinomio g que voy a escoger es: $g = g_1 - h_1g_2 + h_2g_3 = -yz^5 + yz^2 + y$. Notemos que $LT(g) = -yz^5$ no es divisible por $LT(g_i)$*, con* $i = 1, 2, 3$. Por tanto ya hemos encontrado el polinomio que buscabamos.

Observación 3.31. *Sea I un ideal en* $K[x_1, ..., x_n]$*. Vamos a ver que* $G = \{g_1, ..., g_t\} \subseteq I$ *es una base de Groebner de I si y solo si el término líder de cualquier elemento de I es divisible por por uno de los* $LT(g_i)$*.*

Vamos a probarlo por doble implicación:

(i) *Vamos a ver la implicación de izquierda a derecha. Suponemos que* $G = \{g_1, ..., g_t\} \subseteq I$ *es una base de Groebner del ideal* I *y tenemos que llegar a que entonces el término líder de cualquier elemento de* I *es divisible por uno de los* $LT(g_i)$*. Como* G *es base de Groebner entonces para todo elemento* l *de* I *sabemos que* $LT(l) \in \langle LT(g_1), ..., LT(g_t)\rangle$ *y por el Lema 3.18 el* $LT(l) \in \langle LT(g_1), ..., LT(g_t)\rangle$ *si y solo si el* $LT(l)$ *es divisible por algún* $LT(g_i)$*.*

(ii) *Vamos a ver la implicación de derecha a izquierda. Suponemos que para cada* l *en* I*,* $LT(l)$ *es divisible por algún* $LT(g_i)$*. Tenemos que ver que entonces* G *es una base de Groebner de* I *es decir:*

$$\langle LT(g_1), ..., LT(g_t)\rangle = \langle LT(I)\rangle.$$

Para ver esta igualdad comprobamos el contenido $\subseteq$*. Es obvio pues* $g_i \in I$ *y por lo tanto* $LT(g_i) \in LT(I)$ *Ahora sea* $l \in \langle LT(I)\rangle$ *genérico. Como* l *es divisible por algún* $LT(g_i)$ *entonces* $l = h_i \cdot LT(g_i)$ *y por lo tanto* $l \in \langle LT(g_1), ..., LT(g_t)\rangle$*.*

Ejemplo 3.32. *Si usamos el grLex con* $x > y > z$*. ¿Es* $\{l_1 = x^4y^2 - z^5, l_2 = x^3y^3 - 1, l_3 = x^2y^4 - 2zu\}$ *una base de Groebner del ideal generado por esos polinomios?.*

Será una base de Groebner si:

$$\langle LT(l_1), LT(l_2), LT(l_3)\rangle = \langle LT(I)\rangle,$$

es decir, si $\langle LT(I)\rangle = \langle g_1 = x^4y^2, g_2 = x^3y^3, g_3 = x^2y^4\rangle$. Podemos ver lo que nos piden si el LT de cada elemento de I es dividido por algún $LT(l_i)$. Los elementos de I son consecuencias polinómicas de los generadores de I. Cuando tomamos el LT de la combinación $y(x^4y^2 - z^5) - x(x^3y^3 - 1) = -z^5y - x$ obtenemos $-yz^5$ que no es múltiplo de ningún $LT(l_i)$ luego no es base de Groebner.

Observación 3.33. *Sea* I *un ideal principal en* $K[x_1, ..., x_n]$*. Veamos que cualquier subconjunto finito de* I *que contenga a su generador es base de Groebner.*

Como I *es un ideal principal, entonces existe* g *tal que* $I = \langle g\rangle$*. Supongamos que ese suconjunto finito es* $\{g\}$*. En este caso, sí que el subconjunto es base de Groebner ya que para cada elemento de* I *,* f*, se tiene que* $LT(f) \in \langle LT(g)\rangle$*. Supongamos ahora un subconjunto distinto del anterior y finito que contenga a* g*. Como* $I = \langle g\rangle$*, entonces cada* $l \in I$ *es de la forma* $l = h \cdot g$ *con* $h \in K[x_1, ..., x_n]$*. Por tanto el subconjunto finito sería de la forma* $\{g, g \cdot h_1, ..., g \cdot h_t\}$*. Por último para que este subconjunto sea una base de Groebner, todo elemento de I,* f*, ha de cumplir que LT(f) esté en* $\langle LT(g), LT(g \cdot h_1), ..., LT(g \cdot h_t)\rangle$*, lo cual es evidente, puesto que* f *es divisible por* g*.*

Ejemplo 3.34. *Vamos a considerar el ideal* $J = \langle g_1, g_2\rangle$ *con los polinomios* $g_1 = x+z, g_2 = y-z$*. Probaremos que* $\{g_1, g_2\}$ *forman una base de Groebner con el orden monomial Lexicográfico en* $\mathbb{R}[x, y, z]$*. Debemos probar que el término líder de cada elemento no nulo de* J *está en el ideal* $\langle LT(g_1), LT(g_2)\rangle = \langle x, y\rangle$*. Pero por el Lema 3.18 esto es equivalente a probar que el término líder de cualquier elemento no nulo de* J *es divisible por* x *o* y*. Para ello consideremos* $f = A \cdot g_1 + B \cdot g_2 \in J$ *distinto de cero.*

Supongamos que LT(f) no es divisible ni por x *ni por* y*, es decir que es un polinomio únicamente en la variable* z*. Sin embargo* f *se hace cero en el*

subespacio afín $L = \mathbb{V}(x+z, y-z) \subseteq \mathbb{R}^3$ *ya que* $f \in J$. *Notemos que para cualquier* t *real se cumple que* $(x, y, z) = (-t, t, t) \in L$. *El único polinomio que solo contiene la variable* z *y que se hace cero en todos esos puntos es el polinomio igual a cero lo cual es una contradicción. Por tanto* $LT(f)$ *es divisible por* x *o por* y *y por tanto concluimos que* J *es una base de Groebner.*

3.6. Propiedades de las Bases de Groebner.

Vamos a ver que las bases de Groebner son las bases adecuadas de un ideal, en el sentido de que funcionan de modo razonable respecto a nuestro algoritmo de la división.

Proposición 3.35. *Sea* $G = g_1, ..., g_r$ *una base de Groebner de un ideal* $I \subseteq K[x_1, ..., x_n]$ *y sea* $f \in K[x_1, ..., x_n]$. *Entonces hay un único* $r \in K[x_1, ..., x_n]$ *con las siguientes dos propiedades:*

(ii) *Ningún término de* r *es divisible por ningún* $LT(g_1), ..., LT(g_t)$.

(ii) *Existe un* $g \in I$ *tal que* $f = g + r$.

Corolario 3.36. *Sea* $G = \{g_1, ..., g_r\}$ *una base de Groebner de un ideal* $I \subseteq K[x_1, ..., x_n]$ *y sea* $f \in K[x_1, ..., x_n]$. *Entonces* $f \in I$ *si y solo si el resto de la division de* f *por* G *es 0.*

Ejemplo 3.37. *En el Ejemplo 3.34 vimos que* $G = \{x+z, y-z\}$ *es una base de Groebner con el orden monomial Lexicográfico del ideal* $\langle x+z, y-z \rangle$. *Vamos a usar esta base para comprobar la unicidad del resto de la división al cambiar el orden de los divisores. Primero vamos a dividir xy por x+z y a continuación por y-z.*

Por lo tanto $xy = y(x+z) - z(y-z) - z^2$

$A_1 = y$

$x + z$	xy
	$xy + yz$
	$-yz$

$A_1 = -z$

$y - z$	$-yz$
	$-yz + z^2$
	$-z^2$

Ahora vamos a dividir en el orden inverso.

A_1=x

y - z	xy
	xy - xz
	xz

A_1= z

x + z	xz
	$xz + z^2$
	$-z^2$

Por lo tanto $xy = z(x + z) + x(y - z) - z^2$ *y el resto es en ambos casos es* $-z^2$

Definición 3.38. *Resto de la división*
Denotaremos $\overline{f}^F$ *al resto de la división del polinomio f entre la s-tupla* $F = (f_1, ..., f_s)$.

Ejemplo 3.39. *Vamos a realizar la división de* $f = x^5y^3$ *entre* $F = (xy + 1, x^4y^2 - y^2)$. *Tenemos* $F \subseteq \mathbb{R}[x, y]$, *usando el orden monomial Lexicográfico y el algoritmo de la división:*

A_1=xy

A_2=y^2

$x^4y^2 - y^2$	x^5y^3
$xy + 1$	$x^5y^2 - xy^3$
	xy^3
	$xy^3 + y^2$
	$-y^2$

Por lo tanto $x^5y = (y^2) \cdot (xy + 1) + (xy) \cdot (x^4y^2 - y^2) - y^2$. *Finalmente* $\overline{x^5y}^F = -y^2$.

Definición 3.40. *S-polinomio y Mínimo Común Múltiplo*
Sean $f, g \in K[x_1, ..., x_n]$ *tal que* $f, g \neq 0$.

(i) *Si* $multigr(f) = \alpha$ *y* $multigr(g) = \beta$, *entonces sea* $\gamma = (\gamma_1, ..., \gamma_n)$ *donde* $\gamma_i = max\{\alpha_i, \beta_i\}$ *para cada i=1,...,n. Llamaremos Mínimo Común Múltiplo de LM(f) y LM(g)a* x^γ *y lo denotaremos por* $MCM(LM(f), LM(g))$.

(ii) *El S-polinomio de f y g es la combinación*

$$S(f, g) = \frac{x^\gamma}{LT(f)} \cdot f - \frac{x^\gamma}{LT(g)} \cdot g.$$

Ejemplo 3.41. *Sea* $f = x^3y^2 - x^2y^3 + x$ *y* $g = 3x^4y + y^2$ *en* $\mathbb{R}[x, y]$. *Usando el grlex vamos a calcular el S-polinomio de f,g. Tenemos que* $\alpha = multigr(f) = (3, 2)$ *y que* $\beta = multigr(g) = (4, 1)$ *por lo tanto* $\gamma = (4, 2)$. *Sustituyendo en la definición de S-polinomio obtenemos*

$$S(f, g) = \frac{x^4y^2}{x^3y^2} \cdot f - \frac{x^4y^2}{3x^4y} \cdot g = -x^3y^3 + x^2 - \frac{1}{3}y^3.$$

Un S-polinomio surge con la idea de cancelar los términos líder. De hecho, el lema siguiente nos enseñará que toda cancelación de términos líder entre polinomios con el mismo multigrado resulta de este tipo de cancelación.

Lema 3.42. *Sea* $\sum_{i=1}^{s} c_i f_i$ *donde* $c_i \in K$ *y multigr*(f_i)=$\delta \in \mathbb{N}^n$ *para todo i. Si multigr*$(\sum_{i=1}^{s} c_i f_i) < \delta$, *entonces* $\sum_{i=1}^{s} c_i f_i$ *es una combinación lineal, con coeficientes en K, de los S-polinomios* $S(f_j, f_k)$ *para* $1 \leq j, k \leq s$. *Además cada multigr*$(S(f_j, f_k)) < \delta$.

A continuación vamos a ver uno de los resultados clave del trabajo. El lema anterior es muy útil en su demostración.

Teorema 3.43. *Criterio de Buchberger*
Sea I un ideal monomial. Entonces una base $G = \{g_1, ..., g_n\}$ *de I es base de Groebner para I si y solo si para todos los pares* g_j, g_i *con* $i \neq j$, *el resto de dividir* $S(g_j, g_i)$ *por G es cero.*

El criterio de Buchberger dado en el Teorema 3.43 es uno de los resultados más importantes relacionado con las bases de Groebner. Hemos visto algunas de las propiedades de las bases de Groebner, pero hasta ahora era complicado saber si una base de un ideal es una base de Groebner. Sin embargo con este nuevo criterio es sencillo comprobar si una base es de Groebner. En el siguiente apartado veremos que el Criterio de Buchberger nos da un algoritmo para crear bases de Groebner a partir de una base de un ideal dada.

Vamos a presentar una serie de ejemplos sobre los S-polinomios y sobre el Criterio de Buchberger para aclarar los últimos conceptos.

Ejemplo 3.44. *Vamos a calcular con el orden Lexicográfico el S-polinomio de* $f = x^4y - z^2, g = 3xz^2 - y$. *Sustituyendo en la fórmula obtenemos*

$$S(f, g) = \frac{x^4yz^2}{x^4y} \cdot (x^4y - z^2) + \frac{x^4yz^2}{3xz^2} \cdot (3xz^2 - y) = \frac{1}{3}x^3y^2 - z^4.$$

Ejemplo 3.45. *Vamos a ver que* $G = \{-x^2 + y, -x^3 + z\}$ *no es una base de Groebner con el orden monomial Lexicográfico con* $x > y > z$ *para el ideal* $\langle -x^2 + y, -x^3 + z \rangle$. *Para ver si es o no una base de Groebner vamos a usar el Criterio de Buchberger. Para ello calculamos el S-polinomio de* $f = -x^2 + y, g = -x^3 + z$.

$$S_1 = S(f, g) = \frac{x^3}{-x^2} \cdot (-x^2 + y) + \frac{x^3}{-x^3} \cdot (-x^3 + z) = -xy + z.$$

Ahora, para que fuese una base de Groebner el resto de dividir S_1 *entre G tendría que ser 0. Como tanto* g_1 *como* g_2 *tienen monomios líderes de grado mayor que* S_1 *el resto de esa división será* $S_1 \neq 0$ *y por lo tanto no es una base de Groebner.*

3.7. Algoritmo de Buchberger.

En este apartado, daremos un algoritmo para, dada una base de un ideal, poder construir una base de Groebner. En el Corolario 3.26 pudimos ver que todo ideal distinto del vacío poseía una base de Groebner. Pero aún no tenemos una forma de construir bases de Groebner. Dicho esto nos hacemos una pregunta: ¿Dado un ideal $I \subseteq K[x_1, ..., x_n]$, es posible construir una base de Groebner de I?.

Para ver las ideas principales del método que más tarde usaremos, vamos a ver un ejemplo.

Ejemplo 3.46. *Consideremos el anillo de polinomios $\mathbb{R}[x, y]$ con el grlex y sea $I = \langle f_1, f_2\rangle = \langle x^3 - 2xy, x^2y - 2y^2 + x\rangle$. Notemos que $\{f_1, f_2\}$ no forma una base de Groebner de I ya que $LT(S(f_1, f_2)) = -x^2 \notin \langle LT(f_1), LT(f_2)\rangle$.*

Para construir una base de Groebner parece que una idea natural es añadir algún generador extra a la base. ¿ Pero qué generadores deberíamos añadir? Tenemos que $S(f_1, f_2) = -x^2 \in I$ pero $\overline{S(f_1, f_2)}^F = -x^2 \neq 0$ así que vamos a añadir $f_3 = -x^2$ a la nueva versión de la base de I. Vamos a ver, por el Criterio de Buchberger, si $F = (f_1, f_2, f_3)$ es base de Groebner.

$$\overline{S(f_1, f_2)}^F = 0$$

$$S(f_1, f_3) = -2xy \text{ pero } \overline{S(f_1, f_3)}^F \neq 0$$

Así que vamos a añadir $f_4 = -2xy$ a F. Ahora, de la misma forma que antes, vamos a ver si $F = (f_1, f_2, f_3, f_4)$ es base de Groebner.

$$\overline{S(f_1, f_2)}^F = \overline{S(f_1, f_3)}^F = 0$$

$$S(f_1, f_4) = -2xy^2 = yf_4 \text{ por lo tanto}$$

$$\overline{S(f_1, f_4)}^F = 0$$

$$S(f_2, f_3) = -2y^2 + x \text{ pero } \overline{S(f_2, f_3)}^F \neq 0$$

Así que debemos añadir $f_5 = -2y^2 + x$ a F. Aplicando de nuevo el Criterio de Buchberger concluímos que la base de Groebner es

$$F = \{f_1, f_2, f_3, f_4, f_5\} = \{x^3 - 2xy, x^2y - 2y^2 + x, -x^2, -2xy - 2y^2 + x\}.$$

El ejemplo anterior nos sugiere que, en general, deberíamos ir extendiendo nuestra base sucesivamente añadiendo restos distintos de cero de los S-polinomios.

Teorema 3.47. *Algoritmo de Buchberger*
Sea $I = \langle f_1, ..., f_s\rangle \neq \{0\}$ un ideal de polinomios. Entonces una base de Groebner de I puede ser construida en un número finito de pasos por el siguiente algoritmo:

Input: $F = (f_1, ..., f_s)$
Output: Una base de Groebner es $G = (g_1, ..., g_t)$ de I, con $F \subseteq G$

$G := F$
REPEAT

G' := G
FOR cada par $\{p, q\}$, $p \neq q$ *en G' DO*
$S{:=}\overline{S(p,q)}^{G'}$
$IF S \neq 0$ *THEN G':=* $G' \cup \{S\}$
UNTIL G=G'

Demostración.
Comenzamos escribiendo la notación que usaremos durante de demostración. Si $G = \{g_1, ..., g_t\}$ entonces $\langle G\rangle$ y $\langle LT(G)\rangle$ denota los siguientes ideales:

$$\langle G\rangle = \langle g_1, ..., g_t\rangle$$

$$\langle LT(G)\rangle = \langle LT(g_1), ..., LT(g_t)\rangle.$$

Para demostrar el teorema, vamos primero a ver que se cumple $G \subseteq I$ en cada etapa del algoritmo. Esto es cierto inicialmente, cuando agrandamos G, añadimos el resto $S = \overline{S(p,q)}^{G'}$ para p,q∈ G. Así, si $G \subseteq I$, entonces p,q y S(p,q) están en I. Ya que estamos dividiendo por $G' \subseteq I$, tenemos que $G \cup \{S\} \subseteq I$. Notemos que G contiene la base F de I asi que G también es base de I.

El algoritmo termina cuando $G = G'$ lo que significa que $S = \overline{S(p,q)}^{G'} = 0$ para todo p,q de G. Por lo tanto G es una base de Groebner de $I = \langle G\rangle$.

Nos queda probar que el algoritmo termina. Necesitamos ver qué ocurre al finalizar cada paso del bucle principal. El conjunto G consiste en G' (el antiguo G) junto con el resto no nulo de los S-polinomios de los elementos de G'. Entonces

$$\langle LT(G')\rangle \subseteq < LT(G) > \tag{1}$$

ya que $G \subseteq G'$. Además, si $G \neq G'$, veamos que $\langle LT(G')\rangle$ es estrictamente más pequeño $\langle LT(G)\rangle$. Para ver esto suponemos que un resto no nulo r de un S-Polinomio ha sido añadido a G. Ya que r es el resto de la división por G', LT(r) no es divisible por los términos líder de los elementos de G', y así LT(r) $\notin \langle LT(G')\rangle$, además $LT(r) \in \langle LT(G)\rangle$ lo cual prueba lo que queríamos ver.

Por (1), los ideales $\langle LT(G')\rangle$ formados por las sucesivas iteraciones del bucle forman una cadena creciente de ideales en $K[x_1, ..., x_n]$. Así, el Teorema 3.27. implica que después de un número finito de iteraciones la cadena termina. Por lo tanto $\langle LT(G')\rangle = \langle LT(G)\rangle$ debe ocurrir en alguna iteración. Por lo tanto el algoritmo termina en un número finito de pasos.

□

Este algoritmo nos da un método para calcular una base de Groebner pero, al añadir consecutivamente generadores a la base, podríamos acabar con una base demasiado grande. Con los resultados siguientes aprenderemos a eliminar los generadores innecesarios.

Lema 3.48. *Sea G una base de Groebner del ideal de polinomios I. Sea* $p \in G$ *un polinomio tal que* $LT(p) \in \langle LT(G - \{p\})\rangle$. *Entonces* $G - \{p\}$ *también es una base de Groebner de I.*

Definición 3.49. *Base Mínima de Groebner*
Una base mínima de Groebner para un ideal de polinomios I es una base de Groebner G de I que cumple:

(i) $LC(p) = 1$ *para todo* $p \in G$.

(ii) *Para todo* $p \in G$, $LT(p) \notin \langle LT(G - \{p\})\rangle$.

Podemos construir una base mínima de Groebner para un ideal distinto de cero mediante el algoritmo de Buchberger y eliminando los generadores innecesarios que hallamos podido incluir, con el Lema 3.48. Para ilustrar este procedimiento vamos a "reducir"la base de Groebner que habíamos calculado en el Ejemplo 3.46.

Usando el grlex habíamos encontrado la siguiente base de Groebner generada por:

$$\begin{aligned} f_1 &= x^3 - 2xy \\ f_2 &= x^2y - 2y^2 + x \\ f_3 &= -x^2 \\ f_4 &= -2xy \\ f_5 &= -2y^2 + x\ . \end{aligned}$$

Ya que algunos de los coeficientes líder son distintos de 1, lo primero que debemos hacer es multiplicar los generadores por constantes. Una vez hagamos eso, notemos que $LT(f_1) = x^3 = -x \cdot LT(f_3)$, entonces por el Lema 4.48 podemos eliminar f_1 de nuestros generadores.
Igualmente tenemos $LT(f_2) = x^2y = \dfrac{-x}{2} \cdot LT(f_4)$, así que eliminamos f_2 de nuestra base. De esta forma concluimos que una base de Groebner más pequeña es $A = \{\widehat{f_3} = x^2, \widehat{f_4} = xy, \widehat{f_5} = y^2 - \dfrac{x}{2}\}$.

Desafortunadamente, para un ideal dado hay muchas bases mínimas de Groebner. Por ejemplo consideramos para el mismo ideal I que habíamos usado antes, los siguientes generadores para $a \in \mathbb{R}$

$$\widehat{f_3} = x^2 + axy, \quad \widehat{f_4} = xy, \quad \widehat{f_5} = y^2 - \frac{x}{2}.$$

Como $\mathbb{R}$ es infinito, esos polinomios determinan infinitas bases de Groebner distintas. Afortunadamente nosotros podemos elegir una única base que sea mejor que el resto. La siguiente definición nos dirá cual es.

Definición 3.50. *Base Reducida de Groebner*
Una base Reducida de Groebner para un ideal de polinomios I es una base de Groebner G de I tal que:

(i) $LC(p) = 1$ *para todo p en G.*

(ii) *Para todo p en G ningún monomio de p está en* $\langle LT(G - \{p\})\rangle$.

Notar que la única base Reducida en el ejemplo anterior es cuando a=0.

Proposición 3.51. *Sea* $I \neq \{0\}$ *un ideal de polinomios. Entonces para un orden monomial dado, I tiene una única base Reducida de Groebner.*

Para acabar con este apartado, indicaremos brevemente la conexión entre el algoritmo de Buchberger y el algoritmo de eliminación Gaussiana en los sistemas de ecuaciones lineales. Este hecho se debe a que la eliminación Gaussiana es un caso particular del algoritmo que anteriormente hemos visto.

$$\left.\begin{array}{rcl} 3x - 6y - 2z & = & 0 \\ 2x - 4y + 4w & = & 0 \\ x - 2y - z - w & = & 0 \end{array}\right\}$$

Usando operaciones elementales sobre las filas llegamos a la siguiente matriz triangular

$$\begin{pmatrix} 1 & -2 & -1 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Por otro lado si pensamos en el ideal de polinomios

$$I = \langle 3x - 6y - 2z, 2x - 4y + 4w, x - 2y - z - w \rangle \subseteq K[x, y, z, w],$$

usando el algoritmo de Buchberger con el orden monomial Lexicográfico con $x > y > z > w$ conseguimos la siguiente base mínima de Groebner

$G = \langle x-2y-z-w, z+3w \rangle$. Y reduciéndola obtenemos $G = \langle x-2y+2w, z+3w \rangle$

Nos podríamos preguntar ocurrirá siempre, incluso con polinomios no lineales, es decir, si el algoritmo de Buchberger consigue bases Reducidas de Groebner en que los polinomios van teniendo cada vez menos variables. Si nos fijamos esto ha ocurrido también en el Ejemplo 3.46. donde hemos llegado a los polinomios $x^3, xy, y^2 - \frac{x}{2}$. Veremos en la siguiente sección que en efecto esto es así, con lo que las bases de Groebner Reducidas no solo obtienen buenas bases para los ideales que generan los polinomios de las ecuaciones polinómicas si no que además son sistemas equivalentes formados por polinomios que van reduciendo el número de variables, y por tanto, los sistemas son más fáciles de resolver.

Pero antes vamos a ver un ejemplo donde aplicaremos el algoritmo de Buchberger, y luego hallaremos la base mínima descrita anteriormente.

Ejemplo 3.52. *Consideramos el ideal de polinomios*

$$I = \langle g_1 = 3x - 6y - 2z, g_2 = 2x - 4y + 4w, g_3 = x - 2y - z - w \rangle \subseteq \mathbb{R}[x, y, z, w]$$

Usando el orden monomial Lexicográfico con $x > y > z > w$ vamos a calcular una base mínima de Groebner de I.

Antes de nada, vamos a multiplicar por $\frac{1}{2}$ al generador g_2 para facilitar los cálculos. Procedemos calculando los S-polinomios.

$$S(g_1, g_2) = \frac{x}{3x} \cdot (3x - 6y - 2z) - \frac{x}{x} \cdot (x - 2y + 2w) = \frac{-2z}{3} - 2w,$$

$$S(g_1, g_3) = \frac{x}{3x} \cdot (3x - 6y - 2z) - \frac{x}{x} \cdot (x - 2y - z - w) = \frac{z}{3} + w,$$

$$S(g_2, g_3) = \frac{x}{x} \cdot (x - 2y + 2w) - \frac{x}{x} \cdot (x - 2y - z - w) = z + 3w.$$

Como los tres S-polinomios son igual salvo constantes, pues $-3 \cdot S(g_1, g_2) = 6 \cdot S(g_1, g_3) = 2 \cdot S(g_2, g_3)$ solo vamos a dividir $S(g_2, g_3)$ por $\{g_1, g_2, g_3\}$. Debido a que el grado de $S(g_2, g_3)$ es menor que el de cualquier g_i, $\overline{(g_2, g_3)}^I = z + 3w \neq 0$. Tal y como nos dice el algoritmo de Buchberger añadimos a nuestros generadores $g_4 = z + 3w$. Se comprueba entonces que $G = \{g_1, g_2, g_3, g_4\}$, es base de Groebner, pero podemos eliminar g_1 y g_2 debido a que $LT(g_1) = LT(g_2)$ es divisible por $LT(g_3)$.

Por tanto nuestra base de Groebner es $G = \{g_3, g_4\} = \{x-2y-z-w, z+3w\}$ la cual reducida es $G = \{x - 2y + 2w, z + 3w\}$.

4. La Teoría de Eliminación.

Como hemos visto en las secciones anteriores, las soluciones de un sistema de ecuaciones polinómicas son los puntos de la variedad afín determinada por esos polinomios. El sistema original puede ser complicado de resolver, sobre todo si el número de variables es elevado, entonces parece natural la idea de intentar llegar a un sistema de ecuaciones equivalente más sencillo de resolver. Esto es posible gracias a las bases de Groebner, las cuales nos permiten llegar a un sistema más sencillo que el de partida.

En los sistemas de ecuaciones en varias variables con polinomios de grado 1, cuando calculamos la base de Groebner, la cual nos da un sistema equivalente, las ecuaciones obtenidas son las mismas que las que obtenemos con el método de Gauss. En él, a medida que vamos avanzando vamos perdiendo variables hasta llegar finalmente a una ecuación que suele depende de parámetros. En los ejemplos vistos hasta ahora de sistemas polinómicos en varias variables, parecía que pasaba algo parecido al método de Gauss. Si ordenamos los generadores de la base de Groebner por el número de variables que involucran notamos que los generadores van perdiendo variables. Cuando tenemos pocas variables, o una incluso, podremos resolver más cómodamente el sistema equivalente. Aún así, cuando llegamos a polinomios en una variable sólo existen fórmulas para resolver polinomios de grado más pequeño que 5, y si el grado es 5 o superior tendríamos ya que recurrir a métodos numéricos para aproximar la solución. En este capítulo probaremos que el cálculo de bases de Groebner es un método sistemático para eliminar variables, y así veremos que esta eliminación de variables no es una casualidad.

Una vez que calculamos las soluciones del último generador, ¿cómo obtenemos el resto de soluciones?. En el método de Gauss vamos sustituyendo de forma ascendente en el sistema equivalente. Nosotros haremos algo parecido cuando estemos resolviendo un sistema polinómico en varias variables. Este paso lo llamaremos extensión. La extensión de soluciones al resto de ecuaciones será estudiada también en esta sección.

4.1. Los Teoremas de Eliminación y Extensión.

Con el fin de tener una idea intuitiva de que el cálculo de bases de Groebner funciona como una especie de eliminación de variables, vamos a hacer un ejemplo.

Ejemplo 4.1. *Resolveremos el siguiente sistema de ecuaciones.*

$$\left.\begin{array}{rcl} x^2+y+z & = & 1 \\ x+y^2+z & = & 1 \\ x+y+z^2 & = & 1 \end{array}\right\}$$

Tenemos el ideal de polinomios $I = \langle x^2+y+z-1, x+y^2+z-1, x+y+z^2-1\rangle$ *cuya base de Groebner, con el orden Lexicográfico, esta dada por los siguientes polinomios*

$$\begin{aligned} x+y+z^2-1 &= g_1 \\ y^2-y-z^2+z &= g_2 \\ 2yz^2+z^4-z^2 &= g_3 \\ z^6-4z^4+4z^3-z^2 &= g_4 \end{aligned}$$

Ya que en el polinomio g_4 sólo posee la variable z podemos calcular sus soluciones. Como $g_4 = z^2(z-1)^2(z^2+2z-1)$ sus soluciones son 0,1, y $-1 \pm \sqrt{2}$. Sustituyendo esos valores en g_3 y g_2 podemos determinar los valores de la variable y. Sustituyendo esos valores en g_1 obtenemos las cinco posibles soluciones del sistema:

(1,0,0),(0,1,0),(0,0,1),(-1+$\sqrt{2}$, -1+$\sqrt{2}$, -1+$\sqrt{2}$),(-1$\sqrt{2}$, -1$\sqrt{2}$, -1$\sqrt{2}$)

Terminado este ejemplo, parece claro que necesitamos dos pasos para resolver nuestros sistemas:

(i) Eliminación: Necesitamos que entre nuestros polinomios generadores haya algunos con menos variables, y a ser posible con una sóla.

(ii) Extensión: Una vez tengamos las raíces del polinomio con menos variables, tenemos que extender las soluciones para que se verifiquen el resto de condiciones.

Definición 4.2. *Ideal de eliminación*
Dado $I = \langle f_1, ..., f_s \rangle \subseteq K[x_1, ..., x_n]$ el l-ésimo ideal de eliminación I_l es el ideal en $K[x_{l+1}, ..., x_n]$ definido por

$$I_l = I \cap K[x_{l+1}, ..., x_n]$$

Teorema 4.3. *Teorema de la Eliminación*
Sea $I \subseteq K[x_1, ..., x_n]$ un ideal y sea G una base de Groebner del ideal I con respecto al orden monomial Lexicográfico con $x_1 > x_2 > ... > x_n$. Entonces para cada $0 \leq l \leq n$ el conjunto

$$G_l = G \cap K[x_{l+1}, ..., x_n]$$

es una base de Groebner del l-ésimo ideal de eliminación I_l

Demostración. Fijamos un l entre 0 y n. Ya que $G_l \subseteq I_l$, veamos que

$$\langle LT(I_l) \rangle = \langle LT(G_l) \rangle$$

por la definición de base de Groebner. La inclusión $\langle LT(I_l) \rangle \supset \langle LT(G_l) \rangle$ es obvia, para probar $\langle LT(I_l) \rangle \subseteq \langle LT(G_l) \rangle$ debemos probar que LT(f), para un f arbitrario en I_l, es divisible por LT(g) para algún $g \in G_l$.

Para probar esto notar que f también está en I, lo que significa que LT(f) es disible por LT(g) para algún $g \in G$ ya que G es una base de Groebner de I. Ya que $f \in I_l$, esto significa que LT(g) solo posee las variables $x_{l+1,...,x_n}$. Ahora observemos que ya que estamos usando el orden Lexicográfico con $x_1 > ... > x_n$, cualquier monomio que posea alguna variable $x_1, ..., x_l$ es mayor que cualquier monomio en $K[x_{l+1}, ..., x_n]$, así que LT(g)$\in K[x_{l+1}, ..., x_n]$ implica que $g \in K[x_{l+1}, ...x_n]$. Esto demuestra que $g \in G_l$ y termina la demostración. □

Para dar un ejemplo de cómo funciona el Teorema de eliminación vamos a usar el ejemplo 4.1. Considerando $I = \langle x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1 \rangle$ y la base de Groebner con respecto al orden monomial Lexicográfico $G = \{g_1, g_2, g_3, g_4\}$, el Teorema de eliminación nos dice que
$I_1 = I \cap \mathbb{C}[y, z] = \langle y^2 - y - z^2 + z, 2yz^2 + z^4 - z^2, z^6 - 4z^4 + 4z^3 - z^2 \rangle$
$I_2 = \langle z^6 - 4z^4 + 4z^3 - z^2 \rangle$
Así que en la base de Groebner no se han eliminado por casualidad las variables x e y en g_4. El Teorema de eliminación demuestra que una base de Groebner con el orden Lexicográfico va eliminando sucesivamente las variables.

Teorema 4.4. *El Teorema de Extensión*
Sea $I = \langle f_1, ..., f_s \rangle \subseteq \mathbb{C}[x_1, ..., x_n]$ *y sea* I_1 *el primer ideal de eliminación de I. Para cada* $1 \leq i \leq s$*, escribimos* f_i *de la forma*

$$f_i = g_i(x_2, ..., x_n)x_1^{N_i} + \text{ términos en los que } x_1 \text{ tiene grado menor que } N_i,$$

donde $N_i \geq 0$ *y* $0 \neq g_i \in \mathbb{C}[x_2, ..., x_n]$*. Supongamos que tenemos una solución parcial* $(a_2, ..., a_n) \in \mathbb{V}(I_1)$*. Si* $(a_2, ..., a_n) \notin \mathbb{V}(g_1, ..., g_s)$*, entonces existe* $a_1 \in \mathbb{C}$ *tal que* $(a_1, ..., a_n) \in \mathbb{V}(I)$

Para finalizar esta sección vamos a explicar cómo funciona el Teorema de Extensión y a discutir sus consecuencias. La interpretación geométrica será dada en la siguiente sección.

Una primera observación es que el teorema está dado en $\mathbb{C}$. Con el fin de ver la importancia de los complejos, consideramos las siguientes ecuaciones en $\mathbb{R}$

$$x^2 = y,$$

$$x^2 = z.$$

Eliminando x obtenemos y=z, así tenemos la solución parcial (a,a) para todo $a \in \mathbb{R}$. Ya que el coeficiente líder en x de los polinomios $x^2 - z$, $x^2 - y$ nunca se hace 0, el Teorema de la extensión nos garantiza que la solución parcial (a,a) se extiende, pero eso si trabajamos en $\mathbb{C}$, en $\mathbb{R}$ la situación es distinta. Aquí $x^2 = a$ no tiene soluciones reales cuando a es negativo. Así que solo podemos extender aquellas soluciones con $a \geq 0$. Esto demuestra que el Teorema de la extensión falla en $\mathbb{R}$.

Volviendo a la hipótesis $(a_2, ..., a_n) \notin \mathbb{V}(g_1, ..., g_s)$, notemos que los g_i son los coeficientes líder de los f_i. Así, $(a_2, ..., a_n) \notin \mathbb{V}(g_1, ..., g_s)$ nos dice que los coeficientes líder no se hacen cero simultáneamente en la solución parcial. Para ver por qué esta condición es necesaria vamos a echar un vistazo al siguiente ejemplo.

Ejemplo 4.5. *Sean las ecuaciones*

$$xy = 1,$$

$$xz = 1.$$

Notemos que tienen las soluciones parciales $(y, z) = (a, a)$*. La única que no podemos extender es la solución (0,0), que es la solución parcial en la que los coeficientes líder z e y de x se hacen 0. El Teorema de extensión nos dice que el paso de extensión sólo falla cuando los coeficientes líder se hacen cero simultáneamente.*

Es importante añadir que la variedad $\mathbb{V}(g_1, ..., g_s)$ donde los coeficientes líder se hacen cero, depende de la base $\{f_1, ..., f_s\}$ de I. Si cambiamos a distintas bases, es posible que $\mathbb{V}(g_1, ..., g_s)$ cambie. El Teorema de extensión es muy útil cuando los coeficientes líder son constantes, en este caso tenemos el siguiente Corolario.

Corolario 4.6. *Sea* $I = \langle f_1, ..., f_s \rangle \subseteq \mathbb{C}[x_1, ..., x_n]$*, y asumimos que para algún i,* f_i *es de la forma*

$$f_i = cx_1^N + \text{ términos en los que } x_1 \text{ tiene grado menor que } N$$

donde $0 \neq c \in \mathbb{C}$ *y* $N > 0$*. Si* I_1 *es el primer ideal de eliminación de I y* $(a_2, ..., a_n) \in \mathbb{V}(I_1)$*, entonces hay algún* $a_1 \in \mathbb{C}$ *tal que* $(a_1, ..., a_n) \in \mathbb{V}(I)$*.*

Ejemplo 4.7. *Sea el sistema de ecuaciones*

$$\left.\begin{array}{rcl} x^2+2y^2 & = & 2 \\ x^2+xy+y^2 & = & 2 \end{array}\right\}$$

(i) *Si I es el ideal generado por esas ecuaciones, encontraremos bases para $I \cap K[x]$ y $I \cap K[y]$.*

(ii) *Encontraremos todas las soluciones.*

(iii) *Veremos cuales son racionales.*

(iv) *Encontraremos es el cuerpo más pequeño K, tal que contenga todas las soluciones.*

a) La base de Groebner respecto al orden Lexicográfico con $x > y > z$ es dada por los siguientes polinomios

$$g_1 = xy - y^2$$

$$g_2 = x^2 + 2xy - 3$$

Por tanto $I \cap K[y] = I \cap K[x] = \emptyset$

b) Para calcular las soluciones, resuelvo el sistema

$$\left.\begin{array}{rcl} xy-y^2 & = & 0 \\ x^2+2xy-3 & = & 0 \end{array}\right\}$$

obteniendo como soluciones

$$(\pm\sqrt{3}, 0) \qquad ; \qquad (\pm 1, \pm 1)$$

c) Las únicas soluciones racionales son $(\pm 1, \pm 1)$.

d) El cuerpo más pequeño que contiene a todas es $\mathbb{R}$.

Finalmente acabamos esta sección viendo un ejemplo en el que tenemos que aproximar las soluciones.

Ejemplo 4.8. *Consideremos las siguientes ecuaciones*

$$xy = 4$$

$$y^2 = x^3 - 1.$$

Usando el orden Lexicográfico, la base de Groebner es dada por los generadores

$$g_1 = 16x - y^2 - y^4,$$

$$g_2 = y^5 + y^3 - 64,$$

pero si procedemos de la forma usual, descubrimos que g_2 es irreducible sobre $\mathbb{Q}$, es decir, que no tiene raíces racionales. Una opción es calcular las soluciones numéricamente. De esta forma obtenemos:

$$y = 2{,}21363,\ -1{,}78719 \pm 1{,}3984i \text{ ó } 0{,}680372 \pm 2{,}26969i.$$

Estas soluciones pueden ser sustituidas en g_1 y calcular los valores de x.

Como hemos podido apreciar en el último ejemplo, no todos los problemas tienen soluciones racionales. Habrá momentos en los que tengamos que recurrir a métodos numéricos para aproximar las soluciones. También tengamos en cuenta que podemos cambiar el orden monomial e incluso cambiar el orden de las variables para llegar más cómodamente a las soluciones.

4.2. La Geometría en la Eliminación.

En esta sección, daremos una interpretación geométrica de los teoremas vistos en la sección anterior. La idea principal de la eliminación reside en proyectar una variedad en un espacio dimensional más pequeño. También veremos el Teorema de la Clausura, el cual describe la relación entre las soluciones parciales y los ideales de eliminación. En general trabajaremos sobre $\mathbb{C}$.

Definición 4.9. *Proyección de una variedad afín*
Dada la variedad $V = \mathbb{V}(f_1, ..., f_s) \subseteq \mathbb{C}^n$. Para eliminar las primeras l variables, consideramos la proyección

$$\pi_l : \mathbb{C}^n \to \mathbb{C}^{n-l}$$

que envía $(a_1, ..., a_n)$ a $(a_{l+1}, ..., a_n)$. Si aplicamos π_l a $V \subseteq \mathbb{C}^n$ lo denotaremos por $\pi_l(V) \subseteq \mathbb{C}^{n-l}$.

Podemos relacionar $\pi_l(V)$ con el l-ésimo ideal de eliminación de la siguiente forma.

Lema 4.10. *Usando la notación anterior, sea $I_l = \langle f_1, ..., f_s \rangle \cap \mathbb{C}[x_{l+1}, ..., x_n]$ el l-ésimo ideal de eliminación. Entonces, en $\mathbb{C}^{n-l}$ tenemos*

$$\pi_l(V) \subseteq V(I_l)$$

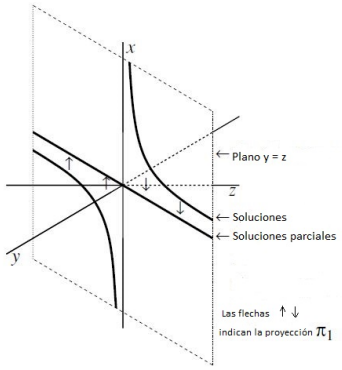
Así, $\pi_l(V)$ consiste exactamente en las soluciones parciales que extendemos para completar las soluciones. Vamos a volver al Ejemplo 4.5 de la sección anterior para ilustrar esto.

Ejemplo 4.11. *Consideremos las ecuaciones*

$$xy = 1,$$

$$xz = 1.$$

En la siguiente imagen podemos ver simultáneamente las soluciones parciales y las soluciones.



Notemos que el primer ideal de eliminación de la base $B = \{-1 + xz, -y + z\}$ es y-z=0 y por tanto $\mathbb{V}(I_1)$ es la recta y=z en el plano yz y que

$$\pi_l(V) = \{(a, a) \in \mathbb{C}^2 \text{ tales que } a \neq 0\}$$

Teorema 4.12. *Dada $V = \mathbb{V}(f_1, ..., f_s) \subseteq \mathbb{C}^n$, sean g_i de la misma forma que en el Teorema de Extensión. Si I_1 es el primer ideal de eliminación de $\langle f_l, ..., f_s \rangle$ entonces tenemos la siguiente igualdad en $\mathbb{C}^{n-1}$*

$$\mathbb{V}(I_1) = \pi_1(V) \cup (\mathbb{V}(g_1, ..., g_s) \cap \mathbb{V}(I_1)),$$

donde π_1 es la proyección en las últimas n-1 variables.

Este teorema nos muestra que $\pi_1(V)$ llena la variedad afín $\mathbb{V}(I_1)$ salvo quizás una parte de $\mathbb{V}(g_1, ..., g_s)$. Desafortunadamente no está claro como de grande es esa parte, y a veces puede ser inmensa.

El siguiente teorema lo enunciamos y probamos para el cuerpo $K = \mathbb{C}$, pero también se puede probar para K un cuerpo algebraicamente cerrado cualquiera.

Teorema 4.13. *Teorema de la Clausura*
Sea $V = \mathbb{V}(f_1, ..., f_s) \subseteq \mathbb{C}^n$ e I_l el l-ésimo ideal de eliminación de $\langle f_1, ..., f_s \rangle$. Entonces:

(i) *$\mathbb{V}(I_l)$ es la variedad afín más pequeña que contiene a $\pi_l(V) \subseteq \mathbb{C}^{n-l}$*

(ii) *Cuando $V \neq \emptyset$, hay una variedad afín $W \subsetneq (I_l)$ tal que $V(I_l) - W \subseteq \pi_l(V)$.*

Demostración. Cuando hablamos de variedad más pequeña en el apartado (i) del teorema, nos referimos a que dicha variedad debe cumplir dos cosas:

- $\pi_l(V) \subseteq \mathbb{V}(I_l)$
- Si Z es otra variedad afín en $\mathbb{C}^{n-1}$ que contiene a $\pi_l(V)$, entonces $\mathbb{V}(I_l) \subseteq Z$.

No probaremos la primera parte del teorema, esto se escapa de los objetivos del trabajo. No obstante podemos encontrarla en el Capítulo 4 de [1].

La segunda parte del teorema dice que a pesar de que $\pi_l(V)$ podría no ser igual a $\mathbb{V}(I_l)$, esta llena "la mayoría"de $\mathbb{V}(I_l)$ en el sentido de que las variables que perdemos están en la variedad afín estrictamente más pequeña. Solo probaremos la parte dos del teorema en el caso $l = 1$. El resto de la demostración es dada en la sección 6 del Capítulo 5 de [1].

La principal herramienta que vamos a usar es la descomposición

$$\mathbb{V}(I_1) = \pi_1(V) \cup (\mathbb{V}(g_1, ..., g_s) \cap \mathbb{V}(I_1)$$

del Teorema de Extesión.

Sea $W = \mathbb{V}(g_1, ..., g_s) \cap \mathbb{V}(I_1)$ y notemos que W es una variedad afín. La descomposición anterior implica que $\mathbb{V}(I_1) - W \subseteq \pi_1(V)$, y de esta forma tenemos el resultado, $W \neq \mathbb{V}(I_1)$. Pero podría pasar que $W = \mathbb{V}(I_1)$.

En este caso, necesitamos cambiar las ecuaciones que definen a V para que W sea más pequeño. La observación clave es que

$$si\ W = \mathbb{V}(I_1),\ entonces\ V = \mathbb{V}(f_1, ..., f_s, g_1, ..., g_s). \tag{2}$$

Vamos a probar esto último. Primero, ya que hemos añadido más ecuaciones, entonces es evidente que $\mathbb{V}(f_1, ..., f_s, g_1, ..., g_s) \subseteq \mathbb{V}(f_1, ..., f_s)$. Para la otra inclusión sea $(a_1, ..., a_n) \in V$. Cada f_i se hace 0 en este punto, y como $(a_2, ..., a_n) \in \pi_1(V) \subseteq \mathbb{V}(I_1) = W$, por lo tanto los g_i también se hacen 0. De

esta forma, $(a_1, ..., a_n) \in \mathbb{V}(f_1, ..., f_s, g_1, ..., g_s)$. Esto último completa la demostración de (2).

Sea $I = \langle f_1, ..., f_s \rangle$ nuestro ideal original, y sea $\widehat{I}$ el ideal $\langle f_1, ..., f_s, g_1, ..., g_s \rangle$. Notemos que I e $\widehat{I}$ podrían ser muy distintos, incluso aunque determinen la misma variedad V. Así el correspondiente ideal de eliminación I_1 e $\widehat{I}_1$ podrían diferir. Sin embargo, ya que $\mathbb{V}(I_1)$ e $\mathbb{V}(\widehat{I}_1)$ son las variedades más pequeñas que contienen a $\pi_1(V)$ (por el apartado (i) del Teorema), entonces $\mathbb{V}(I_1) = \mathbb{V}(\widehat{I}_1)$.

El siguiente paso es encontrar una mejor base para $\widehat{I}$. Primero, recordemos que los $g_i\,'s$ están definidos de la siguiente forma:

$$f_i = g_i(x_2, ..., x_n)x_1^{N_i} + \text{términos con el grado de } x_1 < N_i,$$

donde $N_i \geq 0$ y $g_i \in \mathbb{C}[x_2, ..., x_n]$ no son nulos. Sea el siguiente nuevo conjunto:

$$\widehat{f_i} = f_i - g_i x_1^{N_i}.$$

Para cada i, notemos que $\widehat{f_i}$ es 0 o de grado en x_1 estrictamente más pequeño que f_i. Dejamos como un ejercicio demostrar que

$$\widehat{I} = \langle \widehat{f_1}, ..., \widehat{f_s}, g_1, ..., g_s \rangle.$$

Ahora, aplicamos el Teorema de extensión a $V = \mathbb{V}(\widehat{f_1}, ..., \widehat{f_s}, g_1, ..., g_s)$. Notemos que los coeficientes líder de esos polinomios generadores son distintos, así que obtenemos la siguiente descomposición

$$\mathbb{V}(I_1) = \mathbb{V}(\widehat{I}_1) = \pi_1(V) \cup \widehat{W},$$

donde $\widehat{W}$ consiste en aquellas soluciones parciales en la que los coeficientes líder de $\widehat{f_1}, ..., \widehat{f_s}, g_1, ..., g_s$ se hacen 0.

Antes de seguir con la demostración vamos a realizar un ejemplo para ilustrar cómo $\widehat{W}$ puede ser más pequeño que W. Sea $I = \langle (y-z)x^2 + xy - 1, (y-z)x^2 + xz - 1 \rangle$. Ya que la base de Groebner con el orden Lexicográfico es $\{-y+z, -1+xz\}$, tenemos $I_1 = \langle y-z \rangle$ y $g_1 = g_2 = y-z$, por lo tanto en este caso $W = \mathbb{V}(I_1)$. Entonces es fácil ver que el proceso descrito anteriormente nos da el ideal

$$\widehat{I} = \langle (y-z)x^2 + xy - 1, (y-z)x^2 + xz - 1, y-z \rangle = \langle xy - 1, xz - 1, y - z \rangle.$$

Aplicando el Teorema de Extensión a $\widehat{I}$, nos encontramos con que $\widehat{W}$ consiste en las soluciones parciales en las que las variables z e y se hacen 0, es decir que $\widehat{W} = \{(0, 0)\}$, el cual es estrictamente más pequeño que $W = \mathbb{V}(I_1)$.

Desafortunadamente en el caso general, no hay garantía de que $\widehat{W}$ sea estrictamente más pequeño. Podría ocurrir que $\widehat{W} = \mathbb{V}(I_1)$. En este caso, nosotros repetiríamos el proceso anterior. Si en cualquier etapa del proceso, obtenemos algo estrictamente más pequeño que $\mathbb{V}(I_1)$, habremos acabado.

Nos falta considerar qué es lo que ocurre cuando *siempre* obtenemos $\mathbb{V}(I_1)$. Cada vez que realizamos el proceso anterior, el grado de x_1 de los polinomios generadores disminuye, así que en algún momento el grado en x_1 de los generadores será cero. Esto significa que V podría ser definido por polinomios en $\mathbb{C}[x_2, ..., x_n]$. Así, $(a_2, ..., a_n)$ es una solución parcial, entonces $(a_1, ..., a_n)$ está en V para cualquier $a_1 \in \mathbb{C}$, ya que x_1 no aparece en las ecuaciones que definen al ideal. Por lo tanto toda solución parcial se extiende, lo cual prueba que $\pi_1(V) = \mathbb{V}(I_1)$. En este caso, vemos que el apartado (ii) del teorema es satisfecho cuando $W = \emptyset$. Con esto, el teorema ya está probado.

□

El Teorema de la Clausura nos da una descripción parcial de $\pi_l(V)$ ya que llena $\mathbb{V}(I_l)$ excepto algunos puntos, que viven en una variedad más pequeña que $\mathbb{V}(I_l)$. Pero estos puntos, podrían no llenar toda esa variedad más pequeña.

Corolario 4.14. *Sea $V = \mathbb{V}(f_1, ..., f_s) \subseteq \mathbb{C}^n$, asumimos que para algún i, f_i es de la forma*

$$f_i = cx_1^N + \text{ términos en los que } x_1 \text{ tiene grado menor que } N$$

donde $0 \neq c \in C$ y $N > 0$. Si I_1 es el primer ideal de eliminación, entonces en $\mathbb{C}^{n-1}$

$$\pi_1(V) = \mathbb{V}(I_1),$$

donde π_1 es la proyección en las últimas n-1 variables.

Ejemplo 4.15. *Vamos a ver algún ejemplo sobre esto último. Consideremos $I = \langle f_1, f_2, f_3 \rangle$ donde*

$$f_1 = yx^3 + x^2,$$
$$f_2 = y^3x^2 + y^2,$$
$$f_3 = yx^4 + x^2 + y^2.$$

(i) *Calcularemos una base de Groebner para I y vamos a mostrar que $I_1 = \langle y^2 \rangle$*

(ii) *Probaremos que $\mathbb{V}(I_1) = \mathbb{V}(I_1) \cap \mathbb{V}(g_1, g_2, g_3)$ donde g_i es el coeficiente de la potencia más alta de la variable x en f_i.*

(iii) *Sea $\widehat{I} = \langle f_1, f_2, f_3, g_1, g_2, g_3 \rangle$. Se verá que $I \neq \widehat{I}$ y que $\mathbb{V}(I) = \mathbb{V}(\widehat{I})$*

a) Una base de Groebner para I está generada por los polinomios

$$h_1 = x^2 \qquad h_2 = y^2$$

además, $I_1 = \langle y^2 \rangle$

b) Notemos que $\mathbb{V}(I_1) = \mathbb{V}(I_1) \cap \mathbb{V}(g_1, g_2, g_3)$ es decir que $\mathbb{V}(y^2) = \mathbb{V}(y^2) \cap \mathbb{V}(y, y^3)$.

c) Para ver que $I = \langle f_1, f_2, f_3 \rangle \neq \widehat{I}\langle f_1, f_2, f_3, y, y^3 \rangle$ basta con observar que $y \in \widehat{I}$ pero $y \notin I = \langle h_1, h_2 \rangle = \langle x^2, y^2 \rangle$.

Para probar que $\mathbb{V}(f_1, f_2, f_3) = \mathbb{V}(f_1, f_2, f_3, y, y^3)$, observemos que

$$\mathbb{V}(f_1, f_2, f_3, y) = \mathbb{V}(f_1, f_2, f_3, y, y^3)$$

ya que las soluciones de $y^3 = 0$ e y=0 son las mismas. Ahora bien $\mathbb{V}(f_1, f_2, f_3, y) = \{(0, 0)\} = \mathbb{V}(I)$.

4.3. La Implicitación.

En la sección 2 vimos, que una variedad V puede ser descrita, a veces, usando ecuaciones paramétricas. El procedimiento por el que pasamos de la parametrización a unas ecuaciones que definan la variedad lo llamaremos implicitación. Esta forma de representar a V la llamamos su representación implícita. Pero dada una parametrización: ¿Existe siempre una forma para pasar a su forma implícita?. Dicho esto, nos hacemos dos preguntas. Primero, suponiendo que encontremos la ecuación implícita de la variedad V, ¿llena la parametrización todo

V?. Y segundo, si falta algún punto, ¿cómo los encontramos?. Como veremos las bases de Groebner y el Teorema de extensión serán buenas herramientas para estudiar estas preguntas.

Comenzamos la solución al problema de la implicitación con el caso de la parametrización polinómica. Sea la parametrización:

$$\begin{aligned} x_1 &= f_1(t_1, ..., t_m), \\ &\vdots \\ x_n &= f_n(t_1, ..., t_n). \end{aligned} \tag{3}$$

Aquí, $f_1, ..., f_n$ son polinomios en $K[t_1, ..., t_m]$. Podemos pensar geométricamente en la función

$$F : K^m \longrightarrow K^n$$

definida por

$$F(t_1, ..., t_m) = (f_1(t_1, ..., t_m), ..., f_n(t_1, ..., t_m)).$$

Entonces $F(K^m) \subseteq K^n$ parametrizado por las ecuaciones en (3). Ya que $F(K^m)$ podría no ser una variedad afín, una solución al problema de implicitación sería encontrar la variedad afín más pequeña que contenga a $F(K^m)$. Podemos relacionar la implicitación con la eliminación de variables de la siguiente forma. Las ecuaciones de (3) definen la variedad

$$V = \mathbb{V}(x_1 - f_1, ..., x_n - f_n) \subseteq K^{n+m}.$$

Los puntos de V pueden ser escritos de la siguiente forma

$$(t_1, ..., t_m, f_1(t_1, ..., t_m), ..., f_n(t_1, ..., t_m)),$$

los cuales nos hacen ver que V puede ser la gráfica de la función F. También tenemos dos funciones

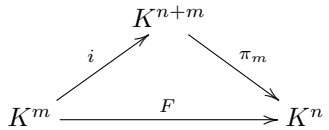
$$i : K^m \longrightarrow K^{n+m}$$
$$\pi_m : K^{n+m} \longrightarrow K^n$$

definidas por

$$i(t_1, ..., t_m) = (t_1, ..., t_m, f_1(t_1, ..., t_m), ..., f_n(t_1, ..., t_m))$$

$$\pi_m(t_1, ..., t_m, x_1, ..., x_n) = (x_1, ..., x_n)$$

Esto nos da el siguiente diagrama:



Notar que F es la composición de funciones $F = \pi_m \circ i$. De esta forma obtenemos que

$$F(K^m) = \pi_m(i(K^m)) = \pi_m(V). \tag{4}$$

Teorema 4.16. *Implicitación Polinómica*
Si K es un cuerpo infinito, sea $F : K^m \longrightarrow K^n$ una función determinada por la parametrización polinómica (3). Sea I el ideal $I = \langle x_1 - f_1, ..., x_n - f_n \rangle \subseteq K[t_1, ..., t_m, x_1, ..., x_n]$ y sea $I_m = I \cap K[x_1, ..., x_n]$ el ideal de eliminación m-ésimo. Entonces $V(I_m)$ es la variedad más pequeña en K^n que contiene $F(K^m)$

Demostración.

Sea $V = \mathbb{V}(I) \subseteq K^{n+m}$. Notemos que V es la gráfica de $F : K^m \to K^n$. Ahora asumimos que $K = \mathbb{C}$. Por (3), tenemos que $F(\mathbb{C}^m) = \pi_m(V)$ y por el Teorema de la Clausura, sabemos que $V(I_m)$ es la variedad más pequeña que contiene a $\pi_m(V)$. Esto prueba el Teorema cuando $K = \mathbb{C}$.

Ahora vamos a suponer K un cuerpo tal que $K \subseteq \mathbb{C}$ y que K hereda las operaciones de $\mathbb{C}$. Este cuerpo siempre contiene a los enteros de $\mathbb{C}$ y ya que es un cuerpo también a $\mathbb{Q}$ y por tanto es infinito. Ya que K puede ser más pequeño que $\mathbb{C}$, no podemos usar el Teorema de la Clausura directamente. Nuestra estrategia será la de ir alternando entre K y $\mathbb{C}$ y usaremos los subíndices $K, \mathbb{C}$ para saber en que cuerpo estamos trabajando. Así $\mathbb{V}_K(I_m)$ es una variedad en K^n, y $\mathbb{V}_{\mathbb{C}}(I_m)$ es una variedad en $\mathbb{C}^n$. Necesitamos probar que $\mathbb{V}_K(I_m)$ es la variedad más pequeña en K^n que contiene a $F(K^m)$.

Por (3) y el Lema 4.10. (que vale también para un cuerpo distinto de los complejos) sabemos que $F(K^m) = \pi_m(\mathbb{V}_K) \subseteq \mathbb{V}_K(I_m)$. Sea $Z_K = \mathbb{V}_K(g_1, ..., g_s) \subseteq K^n$ cualquier variedad en K^n tal que $F(K^m) \subseteq Z_K$. Debemos demostrar que $\mathbb{V}(I_m) \subseteq Z_K$. Comenzamos notando que $g_i = 0$ en Z_K y por lo tanto también en un conjunto más pequeño como es $F(K^m)$. Esto muestra que cada $g_i \circ F$ se hace 0 en todo K^m. Pero g_i es un polinomio en $K[x_1, ..., x_n]$, y $F = (f_1, ..., f_n)$ formado por polinomios en $K[t_1, ..., t_m]$ es decir que $g_i \circ F$ son polinomios en $K[t_1, ..., t_m]$.

De esta forma, los polinomios de la forma $g_i \circ F$ se hacen cero en K^m. Como K es infinito, esto implica que $g_i \circ F$ tiene que ser el polinomio identicamente igual a 0. En particular esto significa que $g_i \circ F$ también se hace cero en $\mathbb{C}^m$ y así los polinomios g_i se hacen cero en $F(\mathbb{C}^m)$. Por lo tanto $Z_{\mathbb{C}} = \mathbb{V}_{\mathbb{C}}(g_1, ..., g_s)$ es una variedad de $\mathbb{C}^n$ que contiene a $F(\mathbb{C}^m)$. Ya que el teorema es cierto para $\mathbb{C}$ entonces $\mathbb{V}_{\mathbb{C}}(I_m) \subseteq Z_{\mathbb{C}}$ *en* $\mathbb{C}^n$. Entonces si nosotros tomamos todas las soluciones que están en K^n entonces $\mathbb{V}_K(I_m) \subseteq Z_K$. Esto prueba que $\mathbb{V}_K(I_m)$ es la variedad más pequeña que contiene a $F(K^m)$.

Finalmente si K es un cuerpo que no está contenido en $\mathbb{C}$, podríamos probar que existe un cuerpo k algebraicamente cerrado contenido en K. Como dijimos en el Teorema de la Clausura, este teorema funciona con cualquier cuerpo algebraicamente cerrado. Y la demostración en ese caso sería similar al procedimiento anterior reemplazando $\mathbb{C}$ por K. □

El último teorema nos da un algoritmo para la implicitación de parametrizaciones polinómicas: Si tenemos $x_i = f_i(t_1, ..., t_m)$ con $f_1, ..., f_m$ polinomios en $K[t_1, ..., t_m]$ consideramos el ideal $I = \langle x_1 - f_1, ..., x_n - f_n \rangle$ y calculamos la base de Groebner con respecto al orden Lexicográfico con $t_i > x_i$. Por el Teorema de eliminación, los elementos de la base que no tienen las variables $t_1, ..., t_m$ forman una base de I_m y por el Teorema 4.16., esos elementos definen la variedad más pequeña en K^n que contiene a la parametrización.

Vamos a ver un ejemplos que muestra esto.

Ejemplo 4.17. *Sea S la superficie parametrizada por las siguientes ecuaciones*

$$x = uv,$$

$$y = u^2,$$

$$z = v^2.$$

(i) *Encontraremos la variedad V más pequeña que contiene a S.*

(ii) *Sobre* $\mathbb{C}$ *mostraremos que S=V.*

(iii) *Nos preguntamos qué pasa para* $K = \mathbb{R}$.

a) Para calcular la variedad V más pequeña que contiene a S, primero calculamos una base de Groebner de $I = \langle x - uv, y - u^2, z - v^2 \rangle$ *usando el orden Lexicográfico con* $u > v > x > y > z$. *La base está generada por los siguientes generadores*

$$g_1 = u^2 - y$$

$$g_2 = uv - x$$

$$g_3 = ux - vy$$

$$g_4 = uz - vx$$

$$g_5 = x^2 - yz.$$

Como podemos ver $I_2 = \langle x^2 - yz \rangle$, *por tanto* $V = \mathbb{V}(I_2)$

b) Vamos a probar que $\mathbb{V}(x^2 - yz) = S$. *Sea un punto de S de la forma* (uv, u^2, v^2) *y lo sustituimos en* g_5 *y obtenemos 0, por tanto* $S \subseteq \mathbb{V}(x^2 - yz)$. *Para el otro contenido vamos a considerar la ecuación* $x^2 - yz$. *Si elijo y,z como parámetros y tomamos* $u = \pm\sqrt{y}$ *ó* $v = \pm\sqrt{z}$ *queremos ver si x=uv. Pero x cumple que* $x^2 = yz$ *es decir, cualquier* $x \in \mathbb{C}$ *que cumple que* $x^2 = yz$ *está en* $\mathbb{V}(x^2 - yz)$, *y notemos que x=uv verifica esto ya que* $x^2 = u^2v^2 = yz$

c)Observamos que en $\mathbb{R}$ *el contenido* $\mathbb{V}(x^2 - yz) \subseteq S$ *no se da. Si y,z son negativos, se pueden encontrar* $(x, y, z) \in \mathbb{V}(x^2 - yz)$ *con* $x^2 = yz$, *pero no están en S ya que y,z no son cuadrados.*

El siguiente paso en nuestro problema de implicitación es ver qué pasa cuando tenemos una parametrización racional. Para ver las dificultades que este tipo de parametrizaciones pueden tener, consideramos la siguiente parametrización racional:

$$x = \frac{u^2}{v}, y = \frac{v^2}{u}, z = u. \tag{5}$$

Es sencillo comprobar que el punto (x,y,z) siempre está en la superficie $x^2y = z^3$. Vamos a ver qué ocurre si quitamos los denominadores y aplicamos el algoritmo de implicitación polinómica. Tendremos el ideal

$$I = \langle vx - u^2 \ , \ uy - v^2 \ , \ z - u \rangle \subseteq K[u, v, x, y, z],$$

Una vez hacemos la base de Groebner obtenemos que $I_2 = \langle z(x^2y - z^3) \rangle$. Esto implica que

$$\mathbb{V}(I_2) = \mathbb{V}(x^2y - z^3) \cup \mathbb{V}(z),$$

y, en particular, $\mathbb{V}(I_2)$ no es la variedad más pequeña que contiene la parametrización. Así que la idea de quitar los denominadores parece no ser la más adecuada.

En general en el caso de parametrizaciones racionales tendremos

$$\begin{aligned} x_1 &= \frac{f_1(t_1,...,t_m)}{g_1(t_1,...,t_m)} \\ &\vdots \\ x_n &= \frac{f_n(t_1,...,t_m)}{g_n(f_1,...,t_m)} \end{aligned} \tag{6}$$

donde $f_1,...,f_n,g_1,...,g_n$ son polinomios en $K[t_1,...,t_m]$. Consideramos $W = \mathbb{V}(g_1,...,g_n) \subseteq K^m$ entonces

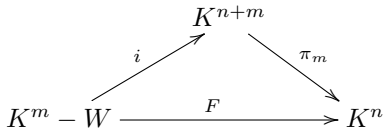
$$F(t_1,...,t_m) = \left(\frac{f_1(t_1,...,t_m)}{g_1(t_1,...,t_m)},...,\frac{f_n(t_1,...,t_m)}{g_n(t_1,...,t_m)}\right)$$

define la aplicación

$$F : K^m - W \longrightarrow K^n$$

Para resolver el problema de la implicitación, necesitamos encontrar la variedad más pequeña de K^n que contenga a $F(K^m - W)$.

Podemos adaptar el diagrama que usamos para parametrizaciones polinómicas para este caso



Sea $I = \langle g_1x_1 - f_1,...,g_nx_n - f_n\rangle$ el ideal que se obtiene "limpiando los denominadores". Tenemos que $i(K^m - W) \subseteq \mathbb{V}(I)$. El problema es que $\mathbb{V}(I)$ podría no ser la variedad más pequeña que contiene a $i(K^m - W)$.

Para evitar esta dificultad vamos a alterar el ideal I añadiendo una dimensión extra para controlar los denominadores. Consideramos el anillo de polinomios $K[y,t_1,...,t_m,x_1,...,x_n]$, sea $g = g_1 \cdot g_2 \cdots g_n$ así que $W = \mathbb{V}(g)$. Consideramos el ideal

$$J = \langle g_1x_1 - f_1,...,g_nx_n - f_n, 1 - gy\rangle \subseteq K[y,t_1,...,t_m,x_1,...,x_n]$$

Notemos que la ecuación $1 - gy = 0$ es decir $1 = gy$ significa que los denominadores $g_1,...,g_n$ nunca pueden hacerse cero en $\mathbb{V}(J)$. Vamos a adaptar el diagrama para estas nuevas situaciones. Sean las aplicaciones

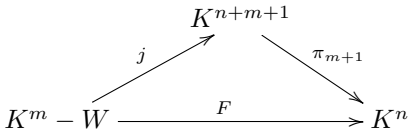
$$\begin{aligned} j : K^m - W &\longrightarrow K^{n+m+1}, \\ \pi_{m+1} : K^{n+m+1} &\longrightarrow K^n \end{aligned}$$

definidas por

$$j(t_1,...,t_m) = \left(\frac{1}{g(t_1,...,t_m)},t_1,...,t_m,\frac{f_1(t_1,...,t_m)}{g_1(t_1,...,t_m)},...,\frac{f_n(t_1,...,t_m)}{g_n(t_1,...,t_m)}\right)$$

$$\pi_{m+1}(y,t_1,...,t_m,x_1,...,x_n) = (x_1,...,x_n)$$

Entonces tenemos el siguiente diagrama



Al igual que antes, tenemos que $F = \pi_{m+1} \circ j$ y además $j(K^m - W) = \mathbb{V}(J)$. Para ver esto notemos que $j(K^m - W) \subseteq \mathbb{V}(J)$ por las definiciones de j y J. Para el otro contenido si $(y, t_1, ..., t_m, x_1, ..., x_n) \in \mathbb{V}(J)$, entonces $gy = 1$ implica que ninguno de los g_i's se hace cero y que $g_i(t_1, ..., t_m)x_1 = f_i(t_1, ..., t_m)$ puede ser escrito de la forma $x_i = \frac{f_i(t_1,...,t_m)}{g_i(t_1,...,t_m)}$.

Debido a que $F = \pi_{m+1} \circ j$ y $j(K^m - W) = \mathbb{V}(J)$ tenemos que

$$F(K^m - W) = \pi_{m+1}(j(K^m - W)) = \pi_{m+1}(\mathbb{V}(J)).$$

Teorema 4.18. *Teorema de la Implicitación Racional*
Si K es un cuerpo infinito, sea $F : K^m - W \longrightarrow K^n$ una función determinada por la parametrización racional (6). Sea J el ideal $J = \langle g_1x_1 - f_1, ..., g_nx_n - f_n, 1 - gy\rangle \subset K[y, t_1, ..., t_m, x_1, ..., x_n]$ donde $g = g_1 \cdot ... \cdot g_n$ y sea $J_{m+1} = J \cap K[x_1, ..., x_n]$ el ideal de eliminación m+1-ésimo. Entonces $V(J_{m+1})$ es la variedad más pequeña en K^n que contiene $F(K^m - W)$.

Demostración. La demostración de este teorema es similar a la del teorema anterior, tendríamos que usar las ecuaciones en (6) en lugar de las ecuaciones en (3). El único punto distinto sería ver que los polinomios que se hacen cero en $K^m - W$ tienen que ser los polinomios igual a 0. □

Concretamente, el Teorema 4.19 nos da un algoritmo para la implicitación de parametrizaciones racionales: si tenemos f_i/g_i polinomios con $f_1, ..., f_n, g_1, ..., g_n \in K[t_1, ..., t_m]$ consideramos $J = \langle g_1x_1 - f_1, ..., g_nx_n - f_n, 1 - gy\rangle$ donde $g = g_1 \cdot ... \cdot g_n$ y la nueva variable y. Calculamos una base de Groebner con respecto al orden Lexicográfico con $y > t_i > x_i$. Los elementos de la base que no tienen las variables $y, t_1, ..., t_m$ definen la variedad en K^n más pequeña que contiene la parametrización.

5. Robótica.

En este capítulo consideraremos una aplicación de los conceptos y técnicas vistos de Geometría algebraica en el área de la Robótica. Desarrollaremos un enfoque sistemático usando variedades, para describir el espacio de posibles configuraciones mecánicas con brazos robóticos, y lo usaremos para resolver los problemas directo e inverso de la Cinemática.

5.1. Descripción Geométrica de Robots.

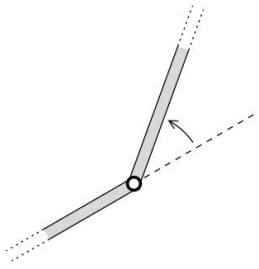
Para tratar geométricamente el espacio de configuraciones de un robot, necesitamos primero simplificar las componentes de nuestros robots y sus propiedades mecánicas. Siempre consideraremos robots construidos a partir de segmentos rígidos, conectados por juntas de distintos tipos. Para simplificar, siempre supondremos que los segmentos están conectados en serie, como las extremidades humanas y consideraremos, por lo general, que un extremo de estos *brazos robóticos* estará en una posición fija. En el otro extremo habrá una *mano* formada por una serie de mecanismos cuyo fin será la de coger objetos o realizar alguna otra tarea. Así, uno de nuestros objetivos será el de intentar describir la posición y orientación de la mano.

Ya que los segmentos de nuestro robot son rígidos, los posibles movimientos de nuestro robot están determinados por el movimiento de las juntas. Actualmente la mayoría de robots están fabricados usando

(i) juntas de revolución planas, y

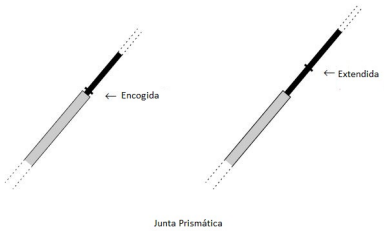
(ii) juntas prismáticas.

Una junta de revolución plana permite la rotación del segmento que une. Asumiremos que los dos segmentos que la junta une, están en el mismo plano, y que todos los movimientos de la junta, mantienen a los segmentos en el mismo plano.



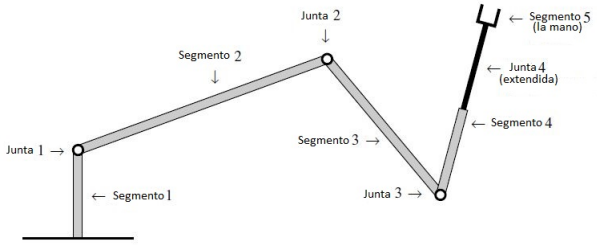
Junta de Revolución

Una junta prismática permite a un segmento del robot trasladarse a lo largo de un eje. El siguiente esquema nos enseña cómo se mueve el segmento con este tipo de juntas.



Si el robot tiene muchas juntas asumiremos, por simplicidad, que todas ellas están en el mismo plano y que los ejes de rotación de las juntas de revolución son perpendiculares al plano. Además los ejes de traslación de las juntas prismáticas viven en el mismo plano. Así, todos los movimientos del robot estarán en el mismo plano. Claro que todo esto nos lleva a clases de robots muy simples, ya que actualmente los robots son capaces de realizar movimientos en 3 dimensiones.

Ejemplo 5.1. *Consideramos el siguiente brazo robótico plano con tres juntas de revolución y una junta prismática. Todos los movimientos del robot estarán en el plano del papel.*



En general la posición de las juntas de revolución entre los segmentos i e i+1 puede ser descrita midiendo el ángulo que forman los segmentos i e i+1. Así la totalidad de posiciones en los que puede estar la junta, pueden ser parametrizados por una circunferencia en el intervalo $[0, 2\pi)$*. En algunos casos la junta no podrá rotar en la totalidad de la circunferencia debido a obstáculos o distintos problemas, en ese caso parametrizaremos un subconjunto de la circunferencia.*

Similarmente, el conjunto de puntos de la junta prismática puede ser especificada dando la longitud de la junta extendida o la longitud del segmento 4 total, es decir la longitud del segmento más la de la junta extendida. De cualquiera de las dos formas anteriores, los puntos a los que la junta prismática podría llegar, pueden ser parametrizados por un intervalo de números reales.

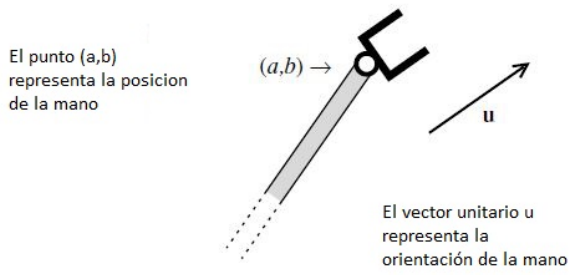
Si los puntos a los que pueden llegar las juntas de nuestro robot pueden ser especificados independientemente, entonces los puntos a los que puede llegar nuestro robot formado por r juntas de revolución y p juntas prismáticas puede ser parametrizado por el producto cartesiano

$$\mathbb{J} = S_1^1 \times \cdots S_r^1 \times I_1 \times \cdots \times I_p,$$

donde cada factor S_i^1 representa una junta de revolución y cada I_j una junta prismática. Llamaremos $\mathbb{J}$ al "conjunto de juntas"de nuestro robot.

Podemos describir los puntos a los que puede llegar la mano del robot plano de la siguiente forma. Fijando un sistema de coordenadas cartesiano podemos

representar los posibles puntos de la mano por los puntos (a,b) de $U \subset \mathbb{R}^2$. Similarmente podemos representar la orientación de la mano por un vector unitario alineado con alguna característica de la mano. De esta forma, las posibles orientaciones de la mano están parametrizadas por los vectores u en $V = S^1$, siendo S^1 una circunferencia. Por ejemplo si la mano está ensamblada a una junta de revolución, tendríamos la siguiente situación



Llamaremos $C = U \times V$ a la configuración espacial de la mano del robot.

Ya que asumimos que los segmentos que forman el robot son rígidos, cada colección de puntos a los que pueden llegar las juntas situará a la mano en una posición única determinada y con una orientación única. Así tenemos la función

$$f : \mathbb{J} \to C$$

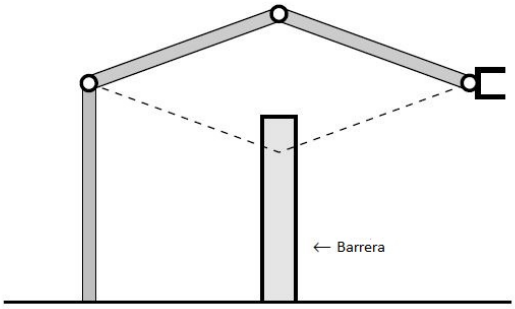
la cual muestra cómo los posibles puntos a los que pueden llegar las juntas producen distintas configuraciones espaciales de la mano.

Los dos problemas básicos que consideraremos pueden ser descritos en términos de la función anterior de la siguiente forma

(i) ¿Podemos dar una descripción explícita o fórmula para f en términos de las juntas de revolución y las longitudes de los segmentos del robot?. A este problema lo llamaremos Problema Directo de la Cinemática.

(ii) Dado $c \in C$ ¿podemos determinar uno de todos los $j \in \mathbb{J}$ tales que $f(j) = c$?. A este problema lo llamaremos Problema Inverso de la Cinemática.

Una característica de los sistemas de ecuaciones no lineales es que puede haber soluciones muy distintas, incluso cuando el número de soluciones es finito. En la subsección 5.3 veremos que esto es cierto para robots con 3 juntas de revolución o más. En la práctica, el robot podría tener obstáculos en ciertos movimientos, esto quiere decir que el robot podría no llegar a ciertos puntos.

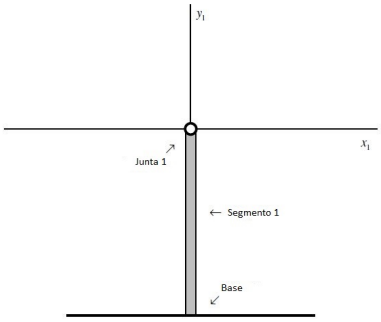


Para saber si es posible llegar a una posición dada c, primero debemos conocer todas las soluciones de $f(j) = c$, y entonces ver cuáles de ellas son compatibles con el entorno en el que está trabajando nuestro robot.

5.2. El Problema Directo de la Cinemática.

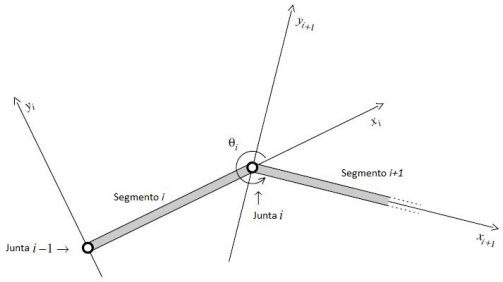
En esta subsección daremos un método para solucionar el Problema Directo de la Cinemática para un brazo de robot dado.

Todos nuestros robots tienen un primer segmento que está fijo en una posición. En otras palabras, que no hay una junta en el extremo inferior del segmento. Con esto, nosotros podemos usar los ejes coordenados usuales para describir la posición y orientación de la mano. El origen de este sistema de coordenadas está situado en la primera junta del brazo del robot la cual esta fija ya que todo el segmento lo está. En el siguiente esquema podemos ver claramente la situación



Además del sistema de referencia global, introducimos un sistema de referencia local para cada junta de revolución para describir las posiciones relativas de los segmentos unidos a esas juntas. Naturalmente, esos sistemas de coordenadas cambiarán conforme cambie la posición de los segmentos.

Para una junta de revolución i, introducimos un sistema de coordenadas (x_{i+1}, y_{i+1}) de la siguiente forma. El origen está situado en la junta i y los ejes (x_{i+1}, y_{i+1}) serán perpendiculares. Observemos que para cada $i \geq 2$, las coordenadas (x_i, y_i) de la junta i son $(l_i, 0)$ donde l_i es la longitud del segmento i.



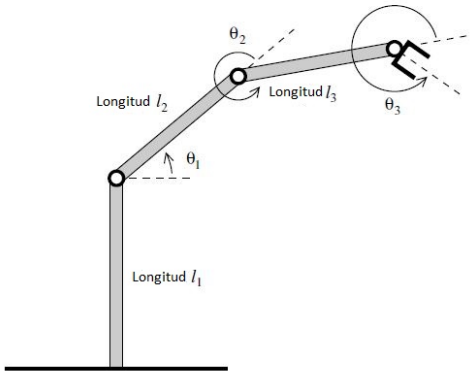
Nuestro primer objetivo es relacionar las coordenadas de un punto en los ejes (x_{i+1}, y_{i+1}) con las coordenadas de ese mismo punto en los ejes (x_i, y_i). Sea σ_i el ángulo que forma el eje x_i con el eje x_{i+1}. De acuerdo al esquema anterior, vemos que si tenemos el punto $q = (a_{i+1}, b_{i+1})$ en el sistema de coordenadas (x_{i+1}, y_{i+1}), para obtener las coordenadas de $q = (a_i, b_i)$ en el sistema de referencia (x_i, y_i) tenemos que rotar usando el ángulo σ_i hasta alinear los ejes x_i e x_{i+1} y entonces trasladar usando el vector $(l_i, 0)$. Así, obtenemos la siguiente relación entre las coordenadas de q en los sistemas de referencia (x_{i+1}, y_{i+1}) y (x_i, y_i)

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} = \begin{pmatrix} \cos(\sigma_i) & -\operatorname{sen}(\sigma_i) \\ \operatorname{sen}(\sigma_i) & \cos(\sigma_i) \end{pmatrix} \cdot \begin{pmatrix} a_{i+1} \\ b_{i+1} \end{pmatrix} + \begin{pmatrix} l_i \\ 0 \end{pmatrix}$$

Esto es usualmente representado con vectores en 3 componentes usando matrices 3x3

$$\begin{pmatrix} a_i \\ b_i \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\sigma_i) & -\operatorname{sen}(\sigma_i) & l_i \\ \operatorname{sen}(\sigma_i) & \cos(\sigma_i) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{i+1} \\ b_{i+1} \\ 1 \end{pmatrix} = A_i \begin{pmatrix} a_{i+1} \\ b_{i+1} \\ 1 \end{pmatrix}$$

Ejemplo 5.2. *Con la notación, vista hasta ahora, consideramos el siguiente brazo robótico con 3 juntas de revolución:*



Pensaremos en la mano del robot como en el segmento 4 que está unido por la junta de revolución 3 al segmento 3. l_i denota la longitud del segmento i. Tenemos

$$A_1 = \begin{pmatrix} \cos(\sigma_1) & -\operatorname{sen}(\sigma_1) & 0 \\ \operatorname{sen}(\sigma_1) & \cos(\sigma_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ya que el origen en el sistema de coordenadas x_2, y_2 está situado en la junta 1. La observación clave es que en el sistema de coordenadas global cualquier punto puede ser obtenido comenzando en el sistema de coordenadas (x_4, y_4) y yendo hacia atrás hasta llegar al sistema (x_1, y_1). Así obtenemos

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = A_1 A_2 A_3 \begin{pmatrix} x_4 \\ y_4 \\ 1 \end{pmatrix}.$$

Usando trigonometría, podemos llegar a

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\sigma_1+\sigma_2+\sigma_3) & -\operatorname{sen}(\sigma_1+\sigma_2+\sigma_3) & l_3\cos(\sigma_1+\sigma_2)+l_2\cos(\sigma_1) \\ \operatorname{sen}(\sigma_1+\sigma_2+\sigma_3) & \cos(\sigma_1+\sigma_2+\sigma_3) & l_3\operatorname{sen}(\sigma_1+\sigma_2)+l_2\operatorname{sen}(\sigma_1) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_4 \\ y_4 \\ 1 \end{pmatrix}.$$

Dado que las coordenadas en (x_4, y_4) son $(0, 0)$, ya que la mano está anclada a la junta 3, obtenemos que las coordenadas de la mano en (x_1, y_1) son:

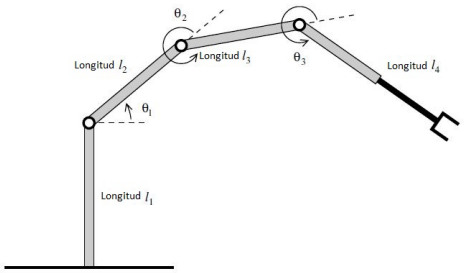
$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} l_3\cos(\sigma_1+\sigma_2)+l_2\cos(\sigma_1) \\ l_3\operatorname{sen}(\sigma_1+\sigma_2)+l_2\operatorname{sen}(\sigma_1) \\ 1 \end{pmatrix},$$

Sabemos que el ángulo formado por el eje x_1 y el eje x_4 es simplemente $\sigma_1+\sigma_2+\sigma_3$. Si combinamos este hecho sobre la orientación de la mano con la fórmula anterior de la posición de la mano podemos describir la función f de la que hablamos en la sección 5.1. Esta función es dada por

$$f(\sigma_1, \sigma_2, \sigma_3) = \begin{pmatrix} l_3\cos(\sigma_1+\sigma_2)+l_2\cos(\sigma_1) \\ l_3\operatorname{sen}(\sigma_1+\sigma_2)+l_2\operatorname{sen}(\sigma_1) \\ \sigma_1+\sigma_2+\sigma_3 \end{pmatrix}.$$

En el caso de que tengamos un número distinto de juntas de revolución, se siguen las mismas ideas. Vamos a realizar un ejemplo más con una junta prismática.

Ejemplo 5.3. *Consideramos un robot plano con 3 segmentos y juntas como los del ejemplo anterior, pero que tiene una junta prismática adicional entre el segmento 4 y el segmento 3. Así que el segmento 4 tendrá una longitud variable y el segmento 5 será la mano.*



El eje de traslación de la junta prismática está situado a lo largo de la dirección del segmento 4. A continuación vamos a describir el robot. Las 3 juntas

de revolución nos permiten los mismos movimientos que las juntas del ejercicio anterior. Pero la junta prismática nos permite cambiar su longitud entre dos valores $l_4 = m_1$ (cuando está encogido) y $l_4 = m_2$ (cuando está totalmente estirado). Al igual que en el ejemplo anterior la configuración espacial de la mano viene dada por

$$g(\sigma_1, \sigma_2, \sigma_3, l_4) = \begin{pmatrix} l_4 \cos(\sigma_1 + \sigma_2 + \sigma_3) + l_3 \cos(\sigma_1 + \sigma_2) + l_2 \cos(\sigma_1) \\ l_4 \operatorname{sen}(\sigma_1 + \sigma_2 + \sigma_3) + l_3 \operatorname{sen}(\sigma_1 + \sigma_2) + l_2 \operatorname{sen}(\sigma_1) \\ \sigma_1 + \sigma_2 + \sigma_3 \end{pmatrix}.$$

Así l_2 y l_3 son constantes, pero tenemos que $l_4 \in [m_1, m_2]$ es una variable más. La orientación de la mano, al igual que en el ejemplo anterior, viene dada por el ángulo $\sigma_1 + \sigma_2 + \sigma_3$ ya que la junta prismática no afecta a la dirección de la mano.

A continuación vamos a discutir cómo las funciones f, g descritas en los ejemplos anteriores puedes ser descritas como funciones polinómicas o racionales. Consideremos la parametrización

$$x = \cos(\sigma)$$

$$y = \operatorname{sen}(\sigma)$$

de la variedad $\mathbb{V}(x^2 + y^2 - 1)$ en el plano. Así podemos escribir las componentes de la función f como

$$c_i = \cos(\sigma_i),$$

$$s_i = \operatorname{sen}(\sigma_i)$$

sujeto a la condición

$$c_i^2 + s_i^2 - 1 = 0$$

para $i = 1, 2, 3$. Notar que la variedad definida por esas tres ecuaciones en $\mathbb{R}^6$ determina el conjunto de juntas $\mathbb{J}$ para este tipo de robot. Geométricamente, esta variedad se corresponde con el producto cartesiano de tres copias del círculo.

Explícitamente obtenemos para f la siguiente expresión en las variables $(c_1, s_1, c_2, s_2, c_3, s_3)$

$$\cos(\sigma_1 + \sigma_2) = \cos(\sigma_1)\cos(\sigma_2) - \operatorname{sen}(\sigma_1)\operatorname{sen}(\sigma_2) = c_1c_2 - s_1s_2$$

$$\operatorname{sen}(\sigma_1 + \sigma_2) = \operatorname{sen}(\sigma_1)\cos(\sigma_2) + \cos(\sigma_1)\operatorname{sen}(\sigma_2) = s_1c_2 + c_1s_2$$

De esta forma las coordenadas de la mano en los ejes (x_1, y_1) están dadas por

$$\begin{pmatrix} l_3(c_1c_2 - s_1s_2) + l_2c_1 \\ l_3(s_1c_2 + c_1s_2) + l_2s_1 \end{pmatrix}.$$

Similarmente obtenemos la forma polinómica de g de la siguiente forma

$$\begin{pmatrix} l_4(c_1(c_2c_3 - s_2s_3) - s_1(c_2s_3 + c_3s_2)) + l_3(c_1c_2 - s_1s_2) + l_2c_1 \\ l_4(s_1(c_2c_3 - s_2s_3) - c_1(c_2s_3 + c_3s_2)) + l_3(s_1c_2 + c_1s_2) + l_2s_1 \end{pmatrix}.$$

Entonces $\mathbb{J}$ es el subconjunto $\mathbb{V} \times [m_1, m_2]$ de la variedad $V \times \mathbb{R}$, donde $V = \mathbb{V}(x_1^2+y_1^2-1, x_2^2+y_2^2-1, x_3^2+y_3^2-1)$. La longitud del segmento 4 será tratado como una variable, así que hemos conseguido ecuaciones que buscábamos en las variables l_4, c_i, s_i.

5.3. El Problema Inverso de la Cinemática.

Para comenzar, vamos a considerar el problema inverso de la cinemática para el brazo robótico del Ejemplo 5.2 con tres juntas de revolución. Dado un punto $(x_1, y_1) = (a, b) \in \mathbb{R}^2$ y una orientación, deseamos determinar si es posible situar la mano del robot en esa orientación. Si es posible, deseamos encontrar todas las combinaciones de movimientos que nos lleven a ese punto. En otras palabras queremos determinar el conjunto de preimágenes de la función $f : \mathbb{J} \to C$, es decir para cada c en C queremos calcular $f^{-1}(c)$.

Es sencillo ver geométricamente que si $l_3 = l_2 = l$ entonces la mano de nuestro robot puede estar situada en cualquier punto del disco cerrado de radio 2l y centrado en la junta 1, es decir el origen del sistema de coordenadas (x_1, y_1). Por otro lado si $l_3 \neq l_2$, entonces las posiciones de la mano llenan un anillo centrado en la junta 1.

Para simplificar nuestro problema inverso de la cinemática, usaremos la siguiente observación para ignorar la orientación de la mano. De esta forma podremos concentrarnos en la posición de la mano, la cual será una función en σ_1 y σ_2. Las posibles formas de colocar la mano en la posición $(x_1, y_1) = (a, b)$ son descritas por el siguiente sistema de ecuaciones polinómicas

$$\left.\begin{array}{lcl} p_1 & : & a = l_3(c_1c_2 - s_1s_2) + l_2c_1 \\ p_2 & : & b = l_3(c_1s_2 + s_1c_2) + l_2s_1 \\ p_3 & : & 0 = c_1^2 + s_1^2 - 1 \\ p_4 & : & 0 = c_2^2 + s_2^2 - 1 \end{array}\right\}$$

para las variables s_1, s_2, c_1, c_2. Para solucionar estas ecuaciones, calcularemos una base de Groebner usando el orden Lexicográfico con las variables ordenadas de la siguiente forma

$$c_2 > s_2 > c_1 > s_1.$$

Nuestras soluciones dependerán de los valores de a, b, l_2, l_3 los cuales aparecen como parámetros en los coeficientes de la base de Groebner generada por los polinomios

$$g_1 = c_2 - \frac{a^2 + b^2 - l_2^2 - l_3^2}{2l_2l_3},$$

$$g_2 = s_2 + \frac{a^2 + b^2}{al_3}s_1 - \frac{a^2b + b^3 + b(l_2^2 - l_3^2)}{2al_2l_3},$$

$$g_3 = c_1 + \frac{b}{a}s_1 - \frac{a^2 + b^2 + l_2^2 - l_3^2}{2al_2},$$

$$g_4 = s_1^2 - \frac{a^2b + b^3 + b(l_2^2 - l_3^2)}{l_2(a^2 + b^2)}s_1 + \frac{(a^2 + b^2)^2 + (l_2^2 - l_3^2)^2 - 2a^2(l_2^2 + l_3^2) + 2b^2(l_2^2 - l_3^2)}{4l_2^2(a^2 + b^2)}.$$

Esta base de Groebner ya está reducida, y es una base para el ideal I generado por los polinomios p_1, p_2, p_3, p_4 en el anillo de polinomios $\mathbb{R}(a, b, l_2, l_3)[s_1, c_1, s_2, c_2]$, observemos que nuestros denominadores depende únicamente de los parámetros a, b, l_2, l_3.

Una primera observación, es que en la práctica, cuando sustituyamos los parámetros a, b, l_2, l_3, necesitaremos que ningún denominador se haga cero, es decir $a \neq 0$ *y* $a^2 + b^2 \neq 0$. Para simplificar las fórmulas, vamos a suponer el caso $l_2 = l_3 = 1$. Obtenemos entonces que

$$g_1 = c_2 - \frac{a^2+b^2-2}{2},$$
$$g_2 = s_2 + \frac{a^2+b^2}{a}s_1 - \frac{a^2b+b^3}{2a},$$
$$g_3 = c_1 + \frac{b}{a}s_1 - \frac{a^2+b^2}{2a},$$
$$g_4 = s_1^2 - bs_1 + \frac{(a^2+b^2)^2-4a^2}{4(a^2+b^2)}.$$

Primero vamos a considerar el caso general $a \neq 0$. Esto implica que $a^2+b^2 \neq 0$ ya que $a, b \in \mathbb{R}$. Resolviendo la ecuación $g_4 = 0$ obtenemos:

$$s_1 = \frac{b}{2} \pm \frac{|a|\sqrt{4-(a^2+b^2)}}{2\sqrt{a^2+b^2}}.$$

Observemos que estas soluciones sólo son reales si $0 < a^2+b^2 \leq 4$, y si $a^2+b^2 = 4$ tenemos una solución doble. La distancia de la junta 1 a la junta 3 es como mucho $l_2 + l_3 = 2$, siendo $\sigma_2 = 0$ en el caso de que sea la distancia igual a 2.

Dado s_1 podemos calcular c_1, s_2, c_2 de los otros polinomios de la base de Groebner. Ya que $a \neq 0$ entonces tenemos un solo valor para las anteriores variables por cada valor de s_1. Por lo tanto los casos cuando $a \neq 0$ ya están hechos.

Ahora vamos a ver los posibles valores de s_1, c_1, s_2, c_2 cuando $a = b = 0$. Geométricamente esto significa que la junta 3 está situada en el origen del sistema de coordenadas (x_1, y_1), es decir el mismo punto en el que está situado la junta 1. La mayoría de nuestros polinomios no están definidos cuando $a = b = 0$, es un caso en el que el modo de resolver nuestro problema falla. Aunque se puede observar en el dibujo que las soluciones son infinitas tomando σ_1 arbitrario y $\sigma_2 = n$.

Finalmente vamos a ver qué ocurre cuando $a = 0,\ b \neq 0$. En este caso también el modo de resolver nuestro problema falla, pero podemos recalcular la base de Groebner habiendo sustituido previamente a por 0. De esta forma obtenemos una nueva base de Broebner dada por los siguientes generadores:

$$g_1' = c_2 - \frac{b^2-2}{2},$$
$$g_2' = s_2 - bc_1,$$
$$g_3' = c_1^2 + \frac{b^2-2}{2},$$
$$g_4' = s_1 - \frac{b}{2}.$$

Notemos que esta nueva base de Groebner tiene diferencias respecto a la anterior. Ahora el último polinomio que tiene como variable s_1 tiene grado 1, en cambio el tercer polinomio tiene grado dos en la variable c_1, por tanto cuando $|b| < 2$ tendremos dos soluciones reales para c_1. Así que cuando $a = 0$ tenemos al menos dos soluciones, que coinciden en la frontera del disco de radio 2.

Vamos a hacer un breve resumen sobre las soluciones de este problema.

(i) Infinitas soluciones cuando $a^2 + b^2 = 0$.

(ii) Dos soluciones distintas cuando $a^2 + b^2 < 4$.

(iii) Una única solución cuando $a^2 + b^2 = 4$.

(iv) Sin soluciones cuando $a^2 + b^2 > 4$.

Conclusiones.

En este trabajo hemos introducido el problema de la resolución de sistemas de ecuaciones polinómicas, mostrando numerosos ejemplos para hacer más comprensible los resultados que hemos visto. Para realizar este estudio hemos empleado numerosas herramientas de carácter básico: álgebra lineal, estructuras algebraicas y geometría.

Como ya hemos dicho en la introducción, la resolución de sistemas polinómicos está presente en infinitud de situaciones. Desde ayudarnos a modelizar diversas situaciones en economía hasta calcular los posibles puntos a los que puede llegar un brazo robótico. No nos olvidemos de la importancia creciente de los robots en la producción industrial y también, por ejemplo, en la investigación espacial.

Personalmente este trabajo me ha enseñado a ver algunas de las aplicaciones de las matemáticas fuera del ámbito académico. Aún nos queda mucho por descubrir, y a medida que la investigación se desarrolla, nuevas aplicaciones surgen de este desarrollo.

Referencias

[1] M.F.Atiyah and I. G. MacDonald, Introduction to Commutative Algebra, Addison- Wesley, Massachusetts, 1969.

[2] T.Becker and V. Weispfenning, *Groebner Bases, Springer-Verlag, New York-Berlin-Heidelberg*, 1993.

[3] B.Buchberger,*Ein algorithmisches kriterium fär die Lösbarkeit eines algebraischen Gleichungssystems* (Alemán), Aequations Math. 4, 1970, 374-383.

[4] B.Buchberger,*Groebner basis: an algorithmic method in polynomial ideal theory*, Multidimensional Systems Theory, editor N.K.Bose, D.Reidel Publishing Company, Dordrecht, 1985, 184-232.

[5] D.Cox, J. Little and D.Oshea, *Ideals, varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, New York, 2007.

[6] D.Eisenbad, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer, New York, 1995.

[7] W.Fulton, *Algebraic Curves*, W.A.Benjamin Juc, New York, 1969.